

## CHAPTER 6 EVALUATION OPTIONS

### I. INTRODUCTION

#### A. Objectives

We present a series of options for evaluating the impacts of welfare reforms on SSA programs. Options are proposed for estimating the impact of non-SSA reforms alone, and for estimating the combined effects of all recent SSA and non-SSA reforms. Several important considerations guided our development of the evaluation options:

- There is a strong consensus among the state and local people we interviewed during our site visits that the conversion of AFDC to TANF and the resulting strict work requirements and time limits had the greatest potential for producing a significant effect on SSA programs. There seems little reason to consider other non-SSA reforms at this time.
- Most of the impacts of non-SSA reforms will be on SSI and any effects on DI are likely to be via concurrent cases only. The only SSA reform that directly impacts DI is the DA&A reform. Even in this instance, 79 percent of the beneficiaries directly affected were SSI recipients, including concurrent recipients (*Lewin, 1998a*). Hence, apart from the evaluation of the effect of DA&A reforms on DI-only cases, it seems sensible to focus evaluation efforts on SSI, with auxiliary analyses of DI where feasible.
- It will be easier to detect the impacts of non-SSA reforms on applications and allowances than on caseloads or payments. Hence, it seems sensible to focus initial evaluation efforts for the effects of non-SSA reforms on applications and allowances.
- It is important to have realistic expectations about the information that a future evaluation can produce. It is unrealistic to expect accurate estimates of the total impacts of all reforms, or of specific non-SSA reforms alone. As was demonstrated by our efforts to model the pre-reform period, it is extremely difficult to account for more than a modest proportion of the factors that are responsible for changes in SSI applications and allowances over time. It is also very difficult to accurately control for the effects of factors such as the economy. Further, two groups, certain children and substance abusers who are parents, are affected by both SSA and non-SSA reforms. This leads us to doubt that any future effort can produce accurate nationwide estimates of the impacts of the non-SSA reforms alone on SSA programs. There is, however, much that can be done to obtain useful information about the interactions between SSA and non-SSA programs, the intersection between the populations they serve, and how they both are changing over time because of program changes and other factors.
- The best way to rigorously evaluate the impacts of non-SSA reforms on SSI is by building on experimental welfare evaluations currently underway. Even though these evaluations will

not produce nationwide estimates of the impacts of reforms, they offer a unique opportunity to establish a causal relationship between specific TANF reforms and SSI outcomes.

## **B. Summary of Options**

The four evaluation options outlined below and described in detail in the remainder of the chapter are designed to achieve the best understanding of the effects of reform that is practical. SSA may choose to pursue some or all of these options. They can be implemented independently of one another or so that the results of various approaches complement and support one another. The first option can provide information for every state as well as the nation as a whole, the second can provide national information and possibly information for some large states, and the final two take advantage of opportunities that are only available in selected states.

### **1. Analysis of SSA Administrative Data**

This option would use SSA administrative data to produce estimates of the impacts of TANF on SSI applications, allowances, caseloads, and benefits, along with collateral estimates of impacts on DI outcomes for those who apply for SSI. It would also produce estimates of the combined impacts of TANF and DA&A reforms on outcomes for those adults who were not SSI recipients at the time the reform legislation was enacted, as well as the combined impacts of TANF and SSI child reforms on outcomes for those children who were not SSI recipients at the time the legislation was enacted. Initial estimates for each post-reform year would be based on age-sex adjusted comparisons of changes in outcomes for target and comparison groups within each state. This would produce time series of estimates for each state, which could be aggregated to obtain national estimates. The reforms might explain any differences observed, although there will inevitably be competing explanations. The state estimates for the impacts of TANF would help SSA detect substantial shifts in SSI applications from, and allowances to, TANF recipients in each state, whether or not they could be definitively attributed to TANF reforms.

SSA could also conduct a pooled time-series analysis of the estimates, to better assess the extent to which TANF reforms and other factors, (e.g., the economy) contributed to the observed shifts. While the marginal value of the pooled time series analysis might be limited, the cost might also be low.

This option also includes a sub-option for evaluating the impact of new restrictions on SSI eligibility for non-citizens.

### **2. Analysis of Census/SSA Matched Data**

The analysis of applications and allowances presented earlier can be extended to produce a second national estimate of the impact of TANF reform on applications and allowances, and auxiliary equations can be developed to generate caseload and benefit estimate. As with Option 1, it will be difficult to disentangle the impacts of TANF from the impacts of other factors on application behavior by TANF recipients. At a minimum, however, it would help SSA detect shifts in participation from TANF to SSI, no matter what the cause. It would also help validate

the national estimates from Option 1. The ability of this option to assess effects in individual states would be very limited.

The matched data can also be used in a relatively simple fashion to estimate the total impacts of the reforms on SSI caseloads in post-reform years, again with the *caveat* that the impacts of “other things” may affect the estimates. They might also be used to improve the analysis of SSA administrative data.

### **3. State Welfare Reform Evaluations**

SSA could select a group of state welfare reform evaluations that are currently underway, and work with existing contractors and states to evaluate the impact of the reforms on SSI. We provide information about the 11 most promising candidates for such “add-on” work, all of which have experimental evaluation designs. These studies would provide methodologically sound estimates of reform impacts in these states. They also would validate the estimates produced under Option 1.

### **4. State Case Studies**

SSA could select states for case studies on the basis of interesting reforms and the availability of data on individuals targeted by TANF reforms. These studies would include both qualitative and quantitative evaluations. The qualitative evaluations would be conducted via structured interviews of key informants. The quantitative evaluations would use SSA data that have been matched to state data from one or more of the following sources:

- Survey and administrative data from the Welfare Leavers studies (14 states or counties);
- Surveys of low-income families in 13 states under the New Federalism project; and/or
- State administrative records.

Such data could be used to produce state estimates that are similar in concept to those described on Option 1, but that take advantage of important information that is not in the SSA data. These would also help validate the Option 1 estimates. We provide information about the states that are the most promising targets for this type of study, and develop a preliminary analysis plan.

In our *Literature Review and Design Report (Lewin, 1998b)*, we included pooled time-series analysis of state application and allowance data, perhaps by age and sex, as a possible primary option. In the options recommended above, pooled time-series analysis is limited to a secondary analysis of the individual state estimates of impacts on applications and allowances derived from SSA administrative data. Although significant improvements can likely be made over the exploratory analysis of the pre-reform period that we presented in Chapter 4, we do not have confidence that the pooled time-series approach can adequately disentangle the effects of reforms from the effects of many other factors that will affect applications and allowances over the period. While cross-state variation in all of these factors offers, in principle, the opportunity to estimate their individual impacts, our ability to measure this variation is too limited for the task at hand.

Instead, we are recommending heavy reliance on within-state analyses that compare changes in outcomes for target and comparison groups. These, too, have limitations, but they also have three distinct advantages over the pooled time-series approach. First, there are opportunities to validate and improve these estimates in some states, provided by the third and fourth options. Second, they potentially control for all factors other than the policy changes that have different impacts on the target and comparison groups, whether or not the factors can be observed. Third, specification errors in one state will not contaminate the estimates in all other states. Further, the evaluator can still use the pooled time-series methodology to help interpret, and refine, the within-state estimates, as we suggest here.

In Section II we discuss the conceptual framework for an evaluation. This is a revision of a discussion that appears in Lewin (1998b), and provides a useful point of reference for the options. The four options listed above are presented in detail in Sections III through VI, respectively. In Section VII we discuss other data sources that we have assessed, some of which may also have value for the evaluation. We discuss the complementarity of the options and their relative strengths and limitations in Section VIII.

## **II. CONCEPTUAL FRAMEWORK**

### **A. Overview**

In this section we discuss each of the following conceptual issues:

- Counterfactual Outcome Series;
- Dynamic Relationships among Primary Outcome Variables;
- DI Outcomes;
- Target and Comparison Groups; and
- Delayed Impacts.

These provide a framework for the evaluation options.

### **B. Counterfactual Outcome Series**

The reforms will have an impact on caseloads, benefit payments and other outcomes every year for the indefinite future, and the size of the impact will vary each year. Hence, it is important to conceptualize the overall objective of the design as estimating and projecting outcome series over the post-implementation period under various policy scenarios.

For each outcome of interest, we would like to compare the series for the outcome under current policy (series A) to estimated series for the following counterfactual scenarios: no policy change (series B); SSA reforms only (series C); and non-SSA reforms only (series D). In *Exhibit 6.1*, we illustrate each of these counterfactual series for a hypothetical outcome.

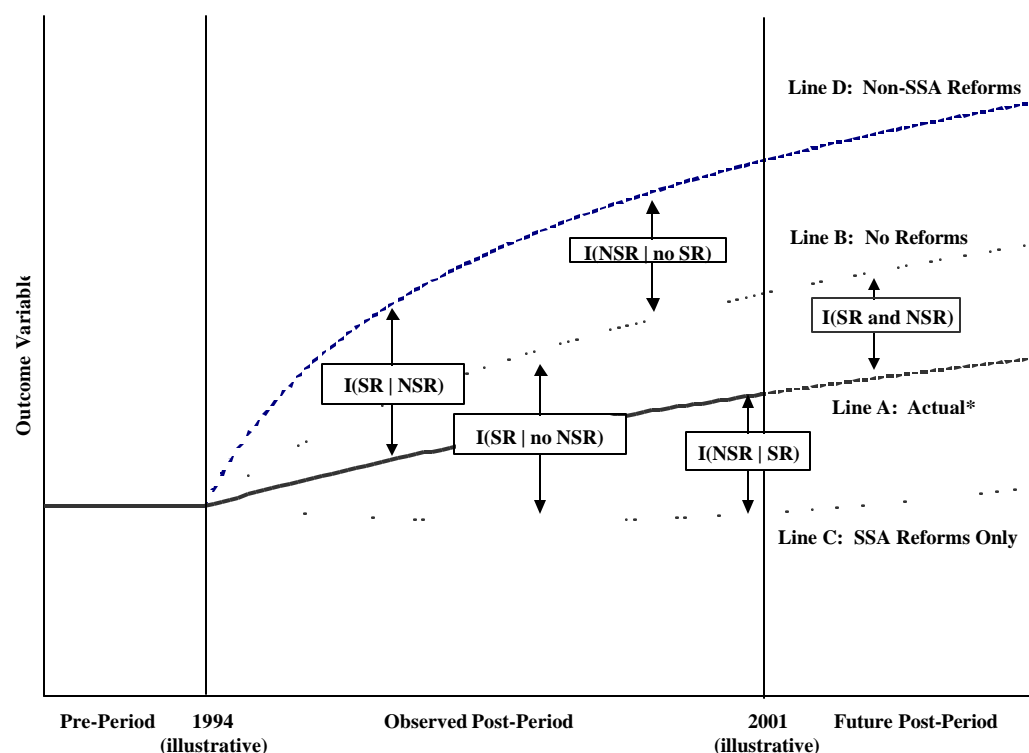
For illustrative purposes, assume the outcome measured in Exhibit 6.1 is SSI caseloads, that impacts on SSI caseloads from SSA reforms are negative, that impacts from non-SSA reforms are positive, and that the magnitude of the impact on SSI caseloads is larger for SSA reforms.

For illustrative purposes, we measure the impact of SSA and non-SSA reforms in comparison to the policies that were in place in July 1994, before the first DA&A policy changes (see *Appendix A*). If the welfare reforms were never implemented, series B would represent future SSI caseload trends based on policies that existed in July 1994. Deviations from series B represent the impacts of SSA and/or non-SSA reforms, and the evaluators task is to estimate the differences between various pairs of these series. The total impact of all policy changes is represented by the difference between series A and B – denoted  $I(\text{SR and NSR})$  in the diagram.

There are two ways to define the impacts of the non-SSA reforms alone. The first is the impact given no SSA reforms,  $I(\text{NSR} \mid \text{no SR})$ , (the difference between series B and D). The second is the impact given the SSA reforms,  $I(\text{NSR} \mid \text{SR})$ , (the difference between series A and C). The latter is expected to be smaller than the former because the two major SSA reforms (concerning DA&A eligibility and child eligibility) likely discourage applications from individuals who might otherwise be induced to apply by the non-SSA reforms, and might also reduce allowance rates for those who apply anyway. Estimating the former is more difficult than the latter because it involves a comparison of two counterfactuals. As a practical matter, the options focus on producing estimates of the impact of non-SSA reforms given the SSA reforms – i.e.,  $I(\text{NSR} \mid \text{SR})$ . This impact is likely the most interesting one to SSA policymakers because it describes how non-SSA policies affect SSA programs given current SSA policy.

An estimate of  $I(\text{NSR} \mid \text{SR})$  along with an estimate of  $I(\text{NSR and SR})$  implies an indirect estimate of the impact of the SSA reforms given no non-SSA reforms:  $I(\text{SR} \mid \text{no NSR}) = I(\text{NSR and SR}) - I(\text{NSR} \mid \text{SR})$ . This may be compared to direct estimates of the impact of SSA reforms, developed by others. Comparability will depend, in part, on whether the direct estimates are conditioned on the non-SSA reforms. If the direct estimate is for  $I(\text{SR} \mid \text{NSR})$ , we would expect them to be smaller in absolute value than the indirect estimate of  $I(\text{SR} \mid \text{no NSR})$ . Also, a direct estimate of  $I(\text{SR} \mid \text{NSR})$  can be used along with an estimate of  $I(\text{SR and NSR})$  to produce an indirect estimate of the impact of the non-SSA reforms given no SSA reforms:  $I(\text{NSR} \mid \text{no SR}) = I(\text{NSR and SR}) - I(\text{SR} \mid \text{NSR})$ . While this estimate may be of less interest to SSA policymakers, it would still be of general interest to know the extent to which the impacts of SSA and non-SSA reforms offset each other, and the information could be of future value to budget and policy analysts.

### Exhibit 6.1 Counterfactual Outcome Series Over the Post Implementation Period Under Various Policy Scenarios\*



\*Line A represents the actual behavior of the outcome variable during the observed post-period and the projected behavior of the outcome variable in the future post-period.

The definition of counterfactual policies requires further discussion. If the objective of the evaluation were to evaluate the impacts of the policy changes induced by the specific federal legislation that is the focus of this project, then the counterfactual policies would be those policies that would be in place if the legislation had not been passed. For SSA reforms, these policies would presumably be the policies that were in place before the legislation. This is not true for non-SSA policies. As discussed in previous chapters, state AFDC programs were changing under federal waivers, some states were implementing other policy reforms in the period before PRWORA was passed, and such reforms likely would have continued had PRWORA not been passed. Determining what these counterfactual policies would have been, and what SSI caseloads would have been under those policies, is virtually an impossible task.

The only practical way to define the counterfactual policies, including state level policies, is to define them as those that were in place in some “base” period, before any of the legislation was passed. To be more accurate, the policy “in place” in a given period should be defined to include recent policy history as well as expected future policies because current behavior depends on both past policies (most recipients in any year became recipients in a previous year) and expected future policies (e.g., anticipated tightening of eligibility for DA&A or child SSI cases).

The definition of post-reform policy is also an issue because the policies of interest changed gradually, and some of the initial changes were partially reversed after a short period. TANF reforms, especially, are being implemented in varying ways and times across states, and it is seems likely that policy changes will continue for some time. The only practical way to define the “post” policy is by the policy that happens to be in place in a given period. That is, for practical purposes the best an evaluation can hope to do is to compare outcomes under current policy in each period to outcomes under the policies in place in some base period. Thus, for instance, the evaluation might be able to answer questions such as “How much of the change in an outcome from 1994 to 1998 can be attributed to changes in welfare policy under TANF?” The evaluation will not be able to answer questions such as “How much different would a 1998 outcome have been from its actual 1998 value had PRWORA not been passed?”

Because the first DA&A legislation was passed in July of 1994, 1993 is the first full calendar year before any reform and the policies in place in that year are a reasonable counterfactual for evaluating the impacts of all reforms. The impact of the early DA&A legislation was likely small in comparison to the legislation passed in 1996 (both the later DA&A legislation and PRWORA). Several states made significant policy changes between 1993 and 1996 (AFDC, GA and others), so for the purpose of studying the impact of TANF, especially, 1995 or even 1996 might be preferred to 1993 for establishing the counterfactual policies. The best option, if feasible, is to try two or three alternative base periods and assess the sensitivity of findings to the choice.

Several of the evaluation options we propose rely on a difference-in-differences (DID) methodology. In general, successful application of DID requires comparison of outcome changes for target and comparison groups from a period that is sufficiently pre-intervention to avoid anticipation effects, to a period that is sufficiently post-intervention to sufficiently capture delayed impacts. Application of the methodology to reforms of interest is problematic because they are phased in and because many of the impacts – especially for TANF – are expected to be substantially delayed. Further, for TANF the phase-in schedule and the impact delays vary substantially across states. Choosing pre and post periods that meet these strict requirements is problematic because the further apart the two periods are, the greater the likelihood that other factors will have differential impacts on the two groups – violating another requirement of DID analysis. Hence, when feasible, we recommend constructing continuous series of DID estimates from before the first of the reforms (no later than 1994) to at least five years after PRWORA (i.e., 2001), to capture what are likely to be phased-in impacts. The possible influence of other factors on these estimates should then be carefully assessed.

### **C. Dynamic Relationships between Primary Outcome Variables**

The outcome variables of primary interest are benefit payments, caseloads, applications, allowances, and terminations. These outcomes are related to one another through a series of dynamic accounting equations (see *Appendix F*). In brief, allowances in a period depend on the number of applications and the share allowed. Caseloads in a period depend on allowances in each past period (“allowance cohorts”) and the share of each allowance cohort that continues to receive payment in the current period. Current benefit payments depend on current program rules and the characteristics of the current caseload. All of these relationships depend on policy history

and the history of external factors that have an impact on applications, allowances and continuation of eligibility.

TANF is not likely to have a measurable impact on SSI outcomes for those who were SSI recipients at the time of PRWORA's passage ("existing recipients"). It is possible, perhaps likely, that TANF might induce some SSI recipients who, in the absence of TANF, would have lost their SSI benefits for various reasons, to stay on the rolls for a longer period. For practical purposes, however, it seems reasonable to ignore any such impact. Hence, we ignore the impact of TANF on existing recipients.

In contrast, the largest immediate impacts of SSA reforms were on terminations of benefits to existing child and DA&A recipients. These reforms are probably also having substantial impacts on the flow into SSI of those who were not recipients at the time the enabling legislation was passed. The evaluation of the total impacts of SSA and non-SSA reforms needs to focus on those who were not existing recipients because all of the reforms affect this group. Impacts of the SSI child and DA&A reforms on existing recipients are being evaluated separately.

#### **D. DI Outcomes**

As argued in the introduction the evaluation of non-SSA reforms should focus on SSI cases only, although DI outcomes for concurrent cases should also be explicitly considered. DI outcomes for SSI applicants can be thought of as SSI applicant characteristics (e.g., percent also applying for DI, percent allowed for DI, etc.).

Over the longer term, TANF reforms may shift payments from SSI to DI through the following mechanism. If the reforms increase the employment of low-income parents in jobs covered by Social Security, then the share of the population that is insured for disability will eventually increase, especially for women. Some of these individuals will qualify for disability payments in the future, perhaps following the later onset of a disability, and the share who will receive at least part of their payment from DI will be higher than it would have been in the absence of the policy change. This will result in a shift in payment costs from general revenues, under SSI, to the DI Trust Fund. Total program costs could increase, but only if the earnings of those affected are high enough to obtain a total payment in excess of the SSI maximum. While it is important for SSA to be aware of this issue, evaluation of the impact would be premature until there is credible evidence that the reforms are increasing covered employment. Evidence from several experimental evaluations discussed later in this chapter may demonstrate that outcome. If so, SSA may want to consider impacts on Trust Fund revenues, insured status of the population, future Trust Fund payments.

#### **E. Target and Comparison groups**

Each of the SSA and non-SSA reforms target specific groups of individuals. The three major reforms of interest have as their target groups:

- SSI and DI DA&A reforms: Recipients and applicants or potential applicants whose drug abuse or alcoholism is material to their disability;



- SSI child eligibility reforms: Recipients and applicants or potential applicants whose disability is based on “maladaptive behavior” or an individual functional assessment;
- SSI benefits for non-citizens: initial reforms made most non-citizens ineligible, but later revisions only made those who immigrated after August 22, 1996 ineligible; and
- Non-SSA reforms: The main reform of interest, TANF, targets members of low-income families with children. Although there are significant exceptions, most such families are one-parent families, and the one-parent is most often a young woman.

Because the reforms are targeted at specific groups, an evaluation that focuses on the target groups will have a better chance of success than one that considers all potential SSI recipients. Further, potential SSI recipients not in target groups can serve as comparison groups in the evaluation – especially if they are similar to those in the target groups in other ways. Use of comparison groups will be very important to a successful evaluation because other significant factors are likely to have an impact on the key outcome variables over the period under study, and because the impacts of those factors will be difficult to measure directly.

To the extent feasible, we incorporate comparison groups in each of the quantitative evaluation options that follow. Options are limited by the data. Because questions are likely to be raised about the validity of any comparison (“natural control”) group, each option should include multiple target – comparison group pairs when feasible.

Because non-SSA reforms vary substantially across states, there is some merit in essentially using states as comparison groups for one another. An important limitation of this approach, however, is that it requires the evaluator to accurately characterize the variation in the nature and timing of reforms across states. Another limitation is that changes in the “other factors” that need to be controlled vary across states. Econometric methods may be used to at least partially control for these, but both our earlier experience and the pre-reform analyses presented in this report indicate that substantial residual differences will remain. While the use of measurable cross-state variation in the non-SSA reforms can enhance the evaluation, we would be skeptical of findings that relied primarily on such variation to measure the impacts of the reforms.

The evaluation needs to compare *changes* in outcome variables for target and comparison groups – essentially using earlier period values for the same groups to control for initial outcome differences. Comparison of post-reform outcomes for target and comparison group cases that do not net out differences in pre-reform outcomes are likely to be misleading because differences at any point in time may be caused by differences between target and comparison groups that are unrelated to policy. Thus, for instance, there would be no value in using the difference between post-period SSI allowances to young women from TANF families to post-period allowances to other young women. There would be value in comparing pre-post changes in allowances to young women from AFDC/TANF families to changes for other young women. Essentially, outcome changes for the comparison group are used to represent the counterfactual for the target group, and the difference between outcome changes for the target and comparison groups is the estimate of the impact of the reform.

There is considerable overlap between the various target and comparison groups. This is most obviously true for children: the SSA child eligibility reforms target children from low-income families who have specific types of conditions, while the AFDC/TANF reforms target low-income families with children. Some parents in the AFDC/TANF target families may also be affected by the DA&A reforms.<sup>267</sup> All of these target groups include both citizens and non-citizens.

As mentioned above, it is not feasible to fully estimate the impact of non-SSA reforms given no-change in SSA policy. The evaluation might, however, make some assessment of this impact by using the intersections of the various target groups. For example, it would be useful to distinguish among: children affected by both the AFDC/TANF reforms and the SSA child reforms; children affected by just the TANF reforms; and children affected by just the SSA child reforms. This will be difficult because the distinguishing features of children in these groups are not cleanly observed in most data. SSA administrative data for existing recipients are somewhat of an exception in this regard. The categories of child SSI recipients that have been eliminated by the SSI child reforms are fairly cleanly defined. Even here, however, there is ambiguity because many children in these categories at the time of the reforms could have qualified in other categories, and some now have. An analogous statement holds for DA&A cases. Worse yet, the post reform data do not clearly identify applicants who would have been allowed under the pre-reform child and DA&A rules.

## **F. Delayed Impacts on Applications and Allowances**

The impacts of reforms on applications and allowances may be delayed substantially. Specific reasons are:

- The many changes in both SSA and non-SSA programs have different implementation dates, so even if effects of individual reforms were immediate, there would be a gradual transition to full effects;
- TANF life-time limits in most states will not be binding for several years, although they may have earlier impacts as families save their time-limited payments for hard times. TANF work requirements are just now starting to bind in most states;
- The currently strong economy may be offering relatively favorable opportunities to many individuals targeted by the reforms who might otherwise seek disability payments. Perhaps just as important, the relatively strong fiscal position of most states may delay efforts to shift TANF recipients onto SSI.

We are concerned that the evaluator may find only small effects of the non-SSA reforms in the immediate future because of these factors, leading policymakers and others to conclusions that neglect what might happen should the economy enter a significant recession. Hence, we favor a design that looks carefully for effects where they are most likely to occur in the short run, and that later evaluates impacts over a much longer period.

---

<sup>267</sup> One study found that 15 percent of adults in AFDC households were substance abusers (Adler, 1993).

## G. Change in Timing of Allowances

Both SSA and non-SSA reforms may accelerate allowances to some individuals who would have received allowances eventually. Even before the non-SSA reforms, substantial numbers of adults and children transitioned from AFDC to SSI every year. Hence, in measuring the impact of TANF of the SSI caseload, or of all reforms on the caseload, we cannot simply assume that increases in allowances induced by the reform translate into future caseload increases after adjustment for attrition. While it will be difficult to identify changes in the timing of allowances directly, it is feasible to produce estimates of caseload impacts for the observed post-observation that implicitly adjusts for such behavior.

## III. ANALYSIS OF SSA ADMINISTRATIVE DATA

### A. Overview

In this section we present an evaluation option that relies almost exclusively on SSA administrative data. This option would use SSA administrative data to:

- Produce state and national estimates of the impacts of TANF on key SSI outcomes (applications, allowances, caseloads and payments);
- Produce state and national estimates of the combined impacts of the TANF, DA&A and child SSI eligibility reforms on key SSI outcomes for those who were not SSI recipients at the time of the reforms;
- Produce state and national collateral estimates of the impacts on DI outcomes for those who apply for SSI; and
- Produce a first-cut analysis of the impacts of non-citizen reforms, just at the national level.

The initial state estimates for the main SSI outcomes would be based on age-sex adjusted comparisons of changes in outcomes for target and comparison groups within each state (“difference-in-differences,” or DID estimates). These would produce time series of estimates for each state (i.e., estimates of impacts for each observed post-reform period). The state estimates could be aggregated to obtain national estimates. We describe this analysis in Section B.

We also describe the first-cut analysis for the impact of non-citizen reforms in Section B, which uses the DID methodology at the national level. Because the initial non-citizen reforms were largely reversed by the BBA, the impacts have become a lower priority for SSA. The first-cut analysis is likely to confirm that effects are small.

While TANF might explain the DID estimates for the impact of TANF, there will inevitably be competing explanations. Even so, the series will be valuable because they will help SSA detect substantial shifts in SSI applications from, and allowances to, TANF recipients in each state.

As stated above, the initial estimates are in the form of time series for each state for the observed post-reform period. SSA might want to construct the same series over the pre-reform period, and

then conduct a pooled time-series analysis of the estimates over the pre- and post-reform periods. This analysis would assess the extent to which TANF reforms contributed to trends in the estimates, and control for changes in the economy and possibly other factors that are controlled for imperfectly in DID analysis. The marginal value of the pooled time series analysis could be limited, but the marginal cost, given that the DID estimates have already been constructed, would likely be low. The analysis would produce refined estimates of the impacts in each state, as well as nationally.

## **B. Difference-in-Differences (DID) Analyses**

### **1. Methodology**

This part of an evaluation would consist of a series of fairly simple “difference-in-differences” (DID) analyses of applications and allowances using SSA administrative data. Most of these analyses would focus on TANF reforms, while others would focus on non-citizen reforms. Over time, DID analyses could be extended to caseload and benefit comparisons.

#### *Overview of Target-Comparison Group Pairs*

We have focused on using information in the Supplemental Security Record (SSR) for the purpose of defining target-comparison group pairs.<sup>268</sup> We have identified the following promising variables for defining target-comparison group pairs: AFDC/TANF income for all applicants at the time of SSI application; immigration status for all applicants at the time of application; parental characteristics for all child applicants at the time of application; family structure information for adult applicants who receive allowances at the time of award; diagnostic information that identifies child applicants who receive allowances and were targeted by the child SSI reforms at the time of award; and similar information that identifies allowed applicants who were targeted by the DA&A reforms. We describe this information and how it may be used later in this section. Most of this information is found in the SSR, but some must be obtained from other sources.

#### *Basic Methodology for Applications and Allowances*

DID analysis compares *changes* (“differences”) in outcomes (applications and allowances) from a period before the policy change (pre-period) to a period after the policy change (post-period) for target-comparison group pairs. The difference between the outcome changes for a target-comparison group pair is the DID estimate for the impact of the policy change on the target group’s outcomes for the post-period. Each DID analysis will produce a series of estimates, comparing outcomes from a series of post-periods to the selected pre-period.

DID analysis is usually applied to percent changes in outcomes, rather than changes in levels, because the pre-period outcomes for the target and comparison groups may be at quite different levels (e.g., applications from parents and non-parents). The assumption is that the outcomes for the two groups would have grown at the same rate in the absence of the policy change. The

---

<sup>268</sup> The SSR is the SSI administrative file that is most accessible for purposes of the evaluation.

estimated change in the level of target group outcomes due to the policy change is computed by multiplying the percentage change estimate by the base period outcome for the target group.<sup>269</sup>

### *Periodicity*

DID analysis could potentially be performed using monthly data. Variation in the number of working days in a month introduces substantial noise into the series, which can be avoided by using moving averages, quarterly data, or even annual data. Use of annual data would minimize the effort required to produce the estimates, but it would be more difficult to relate the timing of policy implementation to the estimated effects on outcomes. In the remainder of this section, we assume that annual data are to be used unless otherwise indicated.

### *Controlling for Age and Sex*

It is important to control for the age and sex composition of the target and comparison groups because:

- the age and sex composition of the groups will differ in most instances;
- SSI applications, allowances and caseloads vary by sex and age;
- the impacts of the policy changes are expected to vary by sex and age; and
- the age distribution of the population changes substantially over the period under study.

This can be accomplished by applying DID analysis to specific age-sex groups. The pattern of findings across age-sex groups is likely to be of interest in itself, and age-sex specific results can be aggregated to obtain estimates of the effects for all age and/or sex groups.<sup>270</sup> We recommend using age groups that are more narrowly defined than those we used in the previous chapter for adults; five-year ranges seem reasonable, with some variation due to critical program ages.

Small cell sizes may become a problem when age ranges are narrowly defined. Estimates for individual age-sex cells may become very unreliable or, worse, undefined. The latter happens when cells are empty for the pre-period, which is used as the base for computing percents. This problem will be especially severe if only a sample of applicants is used, and if quarterly, rather than annual, data are used – issues we return to later in this section. “Smoothing” techniques could be used to address the problem of small cell sizes for continuous variables such as age. These techniques are described in *Exhibit 6.2*.

<sup>269</sup> Mathematically, let  $A_{0t}$  and  $A_{1t}$  be applications from the target group in pre-period 0 and post-period 1, respectively, and let  $A_{0c}$  and  $A_{1c}$  be corresponding values for the comparison group. The percent change in a group's applications from period 0 to 1 is  $\% \Delta A_g = 100 \times (A_{1g} - A_{0g})/A_{0g}$  for  $g = t, c$ . The DID estimate of the growth in applications from the treatment group that is attributed to the policy change, expressed in percentage points, is  $D = \% \Delta A_t - \% \Delta A_c$ . The level of the effect on target group applications is  $L = D \times A_{0t}/100$ .

<sup>270</sup> Continuing the previous footnote, let  $L_{as}$  be the DID estimate of the change in the level of target group applications due to the policy change for those age  $a$  in sex group  $s$ . The aggregated estimate is:

## Exhibit 6.2

### Smoothing Techniques for DID Analysis

The evaluator could use smoothing techniques to address the issue of small cell sizes, if necessary. For instance, to estimate the effects of TANF on SSI outcomes for adult women from a pre-period to a post-period, the evaluator could:

- Fit a continuous distribution to the discrete age distribution for adult female applicants in each of the four relevant applicant groups (pre-period comparison, pre-period target, post-period comparison, and post-period target);<sup>271</sup>
- Use the fitted relationships to “predict” the number of applicants of each year of age in each of the four groups (i.e., produce a smoothed, discrete age distribution of applicants for each group).<sup>272</sup>
- Apply DID to the predictions, for each specific age, to produce estimated effects by individual year of age.
- Add results across ages to obtain estimates for age groups and all women.

The use of wide age ranges to avoid empty cells is a way to smooth applicant age distributions. The approach described can provide a better approximation to the age distribution for an applicant group, avoiding the arbitrarily located steps and flat plateaus of the commonly used method. While the commonly used method is simpler, using it for the purposes of this evaluation could be problematic for the reasons described above.

### State-level Analyses

State-level analyses are warranted whenever feasible, even if the policy change analyzed is a national one, because changes in factors that might be confounded with the effects of policy changes also vary across state. That is, “within state” comparison groups should be used. Within state comparison groups are all the more important because of the fact that the non-SSA policy changes of interest vary across states.

Analysis at the state level allows comparison of the timing of the DID impact estimates to the timing of the implementation of various state reform features, as well as to the timing of changes in confounding factors. The application index graphs that appear in the previous chapter illustrate how that might be done.

A final reason for performing state-level analyses is that the quality of critical data may vary by state. As we discuss later, we are particularly concerned that information on AFDC/TANF receipt at the time of application might be poor in some states.

$$L = \sum_s \sum_a L_{sa} .$$

<sup>271</sup> These could be specific density functions, but it may be more practical and less restrictive to fit a polynomial function.

<sup>272</sup> For each group, the height of the estimated function at a specific age,  $a$ , can be used as the prediction. This should be a very close approximation to the area under the curve in the interval  $(a - 0.5, a + 0.5)$ , which would be the correct way to produce the prediction if the function was a density function. If the curve for a group fits well, the sum of these predictions should be very close to the actual number of applications in the group. There will be discrepancies, however, so it would be sensible to multiply each fitted value by the ratio of actual applications in the group to the sum of the predictions.

SSA might find the fairly simple approach described to this point to be useful for ongoing monitoring of applications and allowances from TANF recipients. We would recommend using 1995 as the pre-period for this analysis, although it would be appropriate to assess the sensitivity of the estimates to use of either 1994 or 1996.

### *Other Applicant Characteristics, Including Those Related to DA&A, Child SSI, and Non-citizen Reforms*

The target and comparison groups may differ by characteristics other than age and sex. SSI applicant characteristics observed in SSA administrative data include race, ethnicity, immigration status, impairment (all in the SSR) and pre-application Social Security earnings (in the Master Earnings File).<sup>273</sup> It is reasonable to think that impacts of TANF and other reforms may vary by these characteristics, just as they are likely to vary by age and sex, and in principal one could further disaggregate the DID analysis. Shrinking numbers of cases in each cell will limit the extent to which this is feasible. It might be best to disaggregate further in the four or five largest states, and not in smaller ones. Immigration status may be especially important in states with large numbers of immigrants, both because of the relatively high growth rate of the immigrant population and the policy changes concerning non-citizens. We present an approach to analysis of these policy changes later in this section.

Attempting to control for impairment differences is problematic because SSA policy changes often change the choices available to the adjudicators. When an applicant has impairments in multiple classes, the adjudicator must choose how to classify the applicant's impairment. Adjudicators are likely to list the impairment that can most easily be demonstrated to meet eligibility requirements, or that comes closest to meeting those requirements. Hence, when the requirements change, the classification of individual cases may change.

It is very important when considering the impact of TANF on SSI outcomes for adults to control for addiction disorders. If the DA&A reforms had differential impacts on the target and comparison groups for the TANF analysis, then the DID estimates of TANF impacts will be biased. Analogously, for analysis of the impact of TANF on child SSI outcomes, it is very important to control for the conditions that were targeted by the SSI child reforms.

It is feasible to control allowance, benefit, and caseload estimates for conditions related to the DA&A and child SSI reforms, but it might not be feasible to do so for application estimates. For the DA&A cases, we recommend using the substance abuse indicator.<sup>274</sup> All SSI recipients designated as DA&A before the DA&A reforms are identified by the codes for the SSR's

<sup>273</sup> The RACE field in the SSR includes Hispanic as a possible response, as well as white, black and several other race or ethnic groups. Applicant tabulations from the 1% SSR Longitudinal File show that race is was not determined in 10.6 percent of cases (Pickett and Scott, 1996). Award tabulations for the same year show that impairment class is missing in only 3.9 percent of cases. Another characteristic that would be substantial interest is education. Education is obtained in the application process, but is not recorded in the SSR. It could potentially be obtained from the 831 File (ED). Education tabulations for SSI applicants from this file that were provided to us in the past by the Office of Disability showed relatively few missing or unknown values.

<sup>274</sup> This variable identifies DA&A cases and also identifies, by separate codes, cases in which substance abuse is known, but not material to disability, and others in which substance abuse is suspected. This information appears in the DRUGM field in the SSR.

substance abuse variable. We expect the variable to identify substance abuse for many other allowed cases even though it was determined to be immaterial to disability.<sup>275</sup>

For children, RAND (1998) has developed a preliminary scheme for identifying cases targeted by the reforms as well as cases in groups to which a substantial share of existing child SSI recipients in the targeted group were reclassified after re-determination.<sup>276</sup> This scheme uses both the primary diagnosis code (PDC) and the Regulation Basis Code (RBC). While the PDC appears on the SSR, the RBC must be obtained from the 831 File. The evaluator should take advantage of RAND's experience in using these codes to classify cases. Any 831 File extracts prepared for the RAND evaluation might also serve the purpose of this evaluation option.<sup>277</sup>

In what follows, we ignore the impact of child SSI reforms on program outcomes for adults. The reforms will, however, reduce the number of SSI children aging into SSI adult status, and may increase applications and allowances for young adults who were not able to obtain benefits as children. Depending on the findings from RAND's evaluation, the evaluator may want to develop identifiers for young adults who are likely to have been in the target group for SSI child reforms when they were children, and control for this characteristic in some analyses.

Whether or not the evaluator controls for any of the other characteristics, it would be interesting to examine the impacts of the reforms on the characteristics of applicants, allowed applicants, and recipients. We return to how this can be done later after we discuss estimation of the impact on caseloads.

### *Estimation of Caseload Effects*

There are two alternative approaches to estimating caseload effects. The first approach applies the DID methodology to administrative data for recently allowed recipients. The second uses a simulation methodology to project the counterfactual caseload for target group applicants that would have been allowed in the post-period, and then compares the counterfactual caseload to the actual caseload for target group cases allowed during the post-period.

The DID methodology described above can be applied fairly straightforwardly to analysis of the impact of TANF on the annual SSI caseload. For the moment we define caseload in a year as the number of individuals eligible for SSI in at least one month during the year ("recipients"). For the first post-reform year, the caseload estimates would be identical to the first-year allowance estimates. For the second post-reform year, the DID methodology would be applied to recipients in the second year who received their allowance in either the first or the second year. For the third post-reform year, the DID methodology would be applied to recipients in the third year who received their allowance in the first, second or third year, and so on.

---

<sup>275</sup> Assessment of substance abuse by adjudicators continues to be important, both for determining materiality to disability and because special restrictions are placed on benefit payments to those with addictive disorders who are able to obtain an allowance for another impairment.

<sup>276</sup> See pp. 78 – 79 of RAND (1998) for a description of their preliminary scheme.

<sup>277</sup> The RAND (1998) design calls for extracting RDC from the 831 File for the entire 10% SSR sample at an unspecified point in time.



This methodology implicitly adjusts for the fact that some target group allowances induced by the reforms in the early post-reform period might be to individuals who would have received allowances later in the period, anyway, because these “advanced” allowances are exactly offset by a decline in later allowances. This is illustrated by means of a stylized example in *Exhibit 6.3*.

One feature of the caseload estimation methodology that needs to be recognized is that when the time period for the estimate advances by one period, the earliest allowance date for those in the estimator’s pre-period caseload moves back one period. Thus, while the caseload in a specific post-period reflects all the policies and other factors that have influenced caseloads in the entire period since the initial policy change, the pre-period caseload reflects these same factors over the same number of years before the policy change. If, for instance, the evaluator estimates the impact on the caseload in 2000, then the pre-period caseload used in the estimates will include individuals who received their allowances as a consequence of both the recession and a variety of policy changes that occurred in the early 1990s. For this reason, the interpretation of the DID caseload estimate will become more problematic as the estimate’s time period advances.

### **EXHIBIT 6.3** **Illustration of the DID Caseload Estimator**

Suppose that:

- In each of the last three years before reform, 500 SSI allowances are made to individuals in the target group, and 500 are made to individuals in the comparison group.
- When an individual receives an SSI allowance, they continue as an SSI recipient for at least two years.
- In the absence of the policy change, annual allowances to the target and control groups would have continued at 500 per year through at least the first three post-reform years.
- The effect of the policy change in the first year is to increase SSI allowances to target group cases by 500 (i.e., to 1,000). The additional allowances include 300 individuals who would have received allowances in the second year anyway, and 200 who would not have received allowances until at least the third post-reform year, if ever.
- In the second year, total allowances to target cases drop to 600, including the 200 remaining cases that would have been allowed in the second year anyway, and 400 cases that would not have been allowed until at least year three, if ever.

In sum, the effect of the policy change on allowances in the first two years is that: 1) 300 allowances to target cases that would have been made in year two are made in year one; and 2) 600 allowances are made that would not have been made until at least year three (200 in year one and 400 in year two).

The DID estimate for allowances and caseloads in the first post-reform year will be 500 (SSI allowances to AFDC/TANF recipients increase from 500 in the pre year to 1,000 in the first post year, and allowances to others are unchanged). The estimate for the second-year caseload increase due to TANF will be 600: 1,000 TANF cases allowed in the first year plus 600 allowed in the second year minus the 1,000 allowed in the two years before reform compared to no-change for non-AFDC/TANF cases. Note that the estimate does not count the 300 allowances moved up from the second post-reform year to the first one as an addition to the second-year caseload.

The alternative approach, simulation of the counterfactual caseload for target group cases that would have been allowed during the post-period, would not rely on the pre-period data. The evaluator would estimate the number of counterfactual allowances during each post-period by

subtracting the DID estimates of allowances induced by the policy change from actual target group allowances, by age and sex. A sample of this number of actual allowances from each year would then be drawn to represent the counterfactual cases, and the actual continuation status in each period of those in the sample would be used to predict the continuation status of the counterfactual cases. The sample could be selected at random within age/sex cells. A refinement would be to use sampling weights that would make sample characteristics within age/sex cells match other predicted characteristics of cases allowed under the counterfactual. Prediction of the latter would require evaluation of the impact of the policy change on those characteristics, which we consider below.

While the simulation approach has the advantage of not relying on the pre-period data, it has the limitation of assuming that, conditional on observed applicant characteristics, continuation patterns are not affected by the policy change – an assumption that is not shared by the DID estimator.

### *Other Program Outcomes*

Other program outcomes of substantial interest include SSI payments, concurrent DI eligibility (for adults), and concurrent DI benefits in each post-allowance period.<sup>278</sup> DI benefits should include dependent benefits, which might substantially add to any DI benefits obtained by a parent. Assuming the analysis is conducted with annual or quarterly data, it will also be of some interest to estimate impacts on the number of months of eligibility during the year. This will be necessary to convert annual or quarterly impacts to mean monthly impacts, in part because published program data are often in monthly units. Other characteristics of recipients could also be thought of as program outcomes, and SSA may find it useful to know how policy changes have affected recipient characteristics (e.g., age, sex, race, impairment, immigration status, and historical earnings).

Two outcome variables of substantial interest could be constructed from historical administrative data: expected lifetime benefit years and expected lifetime benefits. Historical analyses of length-of-stay on SSI that have been conducted previously show that duration is strongly related to age and impairment type.<sup>279</sup> It will be necessary to interpret the results with caution because the reforms could eventually have an impact on length-of-stay and lifetime benefits, holding observable factors constant.

A very simple way to estimate effects on other program outcomes would be to assume that the mean outcomes for all target group cases allowed in a given period (e.g., mean annual SSI payments, percent eligible for DI, etc.) also apply to those allowances induced by the policy change. This could be substantially incorrect, however, and it might be worthwhile to use the DID methodology to obtain refined estimates.

---

<sup>278</sup> We include among concurrent DI cases those who are eligible for SSI before the five-month DI waiting period expires, but who lose SSI eligibility because of their DI benefits once they are DI eligible. This discussion intentionally neglects the impact on DI-only allowances, caseloads and benefit payments, which we have assumed to be small.

<sup>279</sup> See Rupp and Scott (1996, 1998).

The DID methodology can be applied to these other outcomes just as it was to SSI allowances and caseloads. Instead of comparing changes in the number of allowances or the caseload from the target and comparison groups, the evaluator would compare changes in the other outcomes for either those allowed, or those in the caseload and allowed post-reform, from the target and comparison groups. Thus, for instance, to estimate the impact on SSI payments in the first year, the evaluator would compare the percent change in SSI payments made to newly allowed cases from the target group to the percent change in payments made to newly allowed cases from the comparison group. In the second year, the evaluator would compare the same figures for recipients allowed in both the first and the second year.

For immigrants, the analysis needs to consider two additional issues: date of immigration (which is material to the policy changes) and growth in the (legal) immigrant population relative to the native-born population. We will return to these issues later in this section.

### *Using the SSR*

As stated above, the analyses described would primarily use data from the SSR. The SSR is an enormous database, and the evaluator will need to be aware of many issues concerning use of these data. The following discussion of these issues is based on Pickett and Scott (1996).

Each month SSA produces a set of tables based on a 10% SSR sample.<sup>280</sup> These include tables for applications, allowances, recipients, recipient characteristics, and payments. Many are published in the *Social Security Bulletin*. SSA might find it useful to add new tables for some or all of the target and comparison group pairs described below, for the purpose of continually monitoring relative changes in outcomes for these groups.

While convenient, use of the 10% SSR for the evaluation has limitations. One is that cell sizes for application, allowance and caseload counts for target and comparison groups in each state will be quite small in some states every month – especially when divided into age-sex groups. This can be addressed by aggregating the monthly data to quarters or even years. While aggregation to quarters would not result in any substantial loss of information for the evaluation, aggregation to years might because it would be easier to associate changes in outcomes with implementation of new policies using the quarterly data.

A second issue is that caseload counts each month are based on payments actually made during the month, whether or not those payments are based on eligibility for the current month. When allowances are made, eligibility usually begins several months before the allowance date, so persons “paid” in a specific month do not coincide with persons “eligible” in the same month. Differences can be especially large when a policy change or other factor results in an increase or decrease in applications. Changes in allowances as measured by first payment may lag changes in allowances measured by first eligibility date by several months. The relationship between the two allowance series may vary across states. Because the relationship between the timing of outcome changes and timing of policy changes is very important for interpreting the findings,

---

<sup>280</sup> The tabulated application data used in the analysis reported in Chapter 4 are from these samples for the 1991 – 1997 period.

analyses based on eligibility status would likely be more informative than analyses based on paid status.

A second issue is the definition of “application.” Many applications counted in the 10% SSR are “repeat” applications, usually because the first application was denied, but sometimes for administrative reasons. In 1993, about 30 percent of applications were repeat applications. Although many repeat applications represent new attempts to obtain benefits resulting from changes in circumstances or other factors, others are really continuations of previous applications. While some applicants who are denied at the initial determination level will appeal, another strategy to continue pursuit of benefits is to let the appeal period expire and file a new application. Those applicants who do the latter might be counted twice over a period of six months to a year, while those who do the former would only be counted once. This generates noise in the application series, which will make it all the more difficult to identify the impacts of the reforms. While dropping repeat applications from the analysis would be a mistake, it might be better to count only applications that are filed after a substantial period has elapsed since any previous application (e.g., one year).

The last two problems can be addressed using longitudinal data constructed from the SSR. SSA currently constructs a 1% Longitudinal SSR File every six months. This can be used to construct allowance and caseload series based on the eligibility concept, and can also be used to produce application estimates that include only those repeat applications that meet specified conditions. Regrettably, the 1% samples are not large enough to produce state-level DID estimates for many states, or to produce national estimates for immigrants by immigrant cohort.

SSA might find it useful to pursue a strategy that:

- Uses the monthly 10% SSR to follow outcomes for the various target and comparison groups for a few years,
- Builds a special purpose longitudinal SSR file for the purposes of a later evaluation, if the analysis of the 10% SSR indicates that further evaluation would be worthwhile.

SSA already has developed the methodology and software to produce longitudinal records from the SSR for any identified individual. The first step in building the special purpose file would be to identify a large set of potential target and comparison group cases from the full SSR or other administrative data files – enough in each state to produce sufficiently reliable state estimates of the desired statistics. The next step would be to create the longitudinal records for the selected cases.

## **2. DID Estimators for the Impacts of TANF**

### *Estimators Based on AFDC/TANF Receipt at Time of Application (Adults and Children)*

The potentially most useful DID analysis of outcomes to be performed with SSA administrative data alone would distinguish between outcomes for those applicants who are in AFDC/TANF families at the time they apply from those who are not. This analysis can be applied to child

cases as well as adults, and can produce estimates for all of the program outcomes of interest, including collateral estimates for DI outcomes.

A flag on the SSR for “cash assistance based on need” in the month in which the applicant applied for SSI indicates whether the applicant is in an AFDC or TANF family at that time. It is our understanding that the flag, which is called IUETYP in the SSR, is supposed to be coded as “F” in such cases.<sup>281</sup>

This information is ascertained by a SSA Claims Representative (CR) when the SSI application is taken at a field office.<sup>282</sup> The information may be more reliable for allowances than for applications because it is material to the initial SSI payment.<sup>283</sup> It is also possible that some applicants identified as receiving AFDC income were, in fact, receiving general assistance.<sup>284</sup> We have learned through our site visits that at least some field offices try to verify the information with local welfare agencies. This practice may not be uniform and may depend on the cooperation of the local welfare agency. Thus, validity of the information may vary by locality or state, and actual analysis of the data may show that it is not sufficiently reliable to be used for the evaluation’s purposes in some states.

We understand that SSA has used this variable to produce national estimates of the percent of current SSI beneficiaries who are former AFDC recipients. SSA estimates that over 1.4 million of 6.5 million SSI recipients in January 1998 had received AFDC income at the time of application.<sup>285</sup>

The analysis could follow the DID methodology described earlier in this section in a straightforward manner:

- Computation of sex-specific DID estimates for narrowly defined age groups in each state, starting in a specified period. For adults, the evaluator should control for addictive disorders. For children, the evaluator should control for PDC/RBC groups likely to be affected by the SSI child reforms;

<sup>281</sup> This discussion is based on information provided to us by SSA Field Office staff during site visits and Mary Barbour of ORES.

<sup>282</sup> The question is asked at the first interview, when the CR usually completes the “short-form” version of the SSI application, SSA-8001-F5 (7-90) (Question 18), but some times completes the “long-form,” SSA-8000-BK (5-90, especially if an allowance appears likely. The long form must be completed for allowed cases. The question on the short-form is somewhat ambiguous about identifying AFDC/TANF income. It tells the applicant to “List all income received or expected to be received since the first moment of the filing date month. List cash, checks, and direct payments to bank accounts you (your spouse/parents) received or expect to receive. Include income from wages, ....., assistance based on need, .....” The applicant is required to enumerate income by source. The long form asks (Question 31a) “Since the first moment of the filing date month, have you received or do you expect to receive income in the next 14 months from any of the following? A list of income types is provided, including “Aid to Families with Dependent Children” as a line item.

<sup>283</sup> It may also be that the information for allowed cases is more likely to be based on the long-form question, which is more explicit about AFDC/TANF income.

<sup>284</sup> This would most likely occur if income identification is based on the short form because the short-form question does not explicitly distinguish between AFDC/TANF income and other cash assistance based on need.

<sup>285</sup> We are indebted to Charles Scott for providing this information. The estimate appears in an appendix to the *SSI Annual Report* (see SSA, 1998).

- Aggregation to the state level for comparison of the DID series to the timing of policy changes and other events. Separate child, adult female, and adult male series would be warranted. Examination of series for “young” and “older” adults, perhaps split at age 45, is also warranted, to verify that the estimated effects are more substantial for young adults; and
- Aggregation across states to obtain national series.

There are two important limitations of this analysis. One is the previously mentioned possibility that the data are of poor quality, at least in some states. Another is that this method will classify individuals who have received AFDC/TANF, but who left before applying, as comparison cases. If some TANF recipients apply only after they lose benefits because of sanctions or expiration of the lifetime limit, they will be classified as comparison cases. Further, some applicants who in the past might have obtained AFDC before applying for SSI might be deterred from applying for TANF under the new policy, again resulting in classification in the comparison group. Thus, there is some danger that this method could lead to an underestimate of the impact of TANF on applications and allowances.

Whether or not this is a major problem can be ascertained through complementary analyses. One is an analysis that uses state administrative data linked to SSA records, which we return to later in the chapter. The second is DID analysis of SSI allowances for adults using target-comparison group pairs that are based on parenthood status at the time of application. For children, DID analysis based on parent characteristics can serve this purpose. We consider these alternatives next.

### *Estimators Based on Parenthood Status (Adults)*

When an SSI allowance is made, an SSA Claims Representative obtains information about relatives in the family. This information is not available for denied applicants, so it can only be used to produce estimates of impacts on SSI allowances, caseloads and benefits.<sup>286</sup>

The evaluator might use this information to classify adult applicants into three “parenthood status” groups: (1) married parent of a child under 18 who is living with the applicant (“married parent”); (2) unmarried parent of child under 18 who is living with the applicant (“unmarried parent”); and (3) no children under 18 living with the applicant (“non-parent”).<sup>287</sup>

This information is not, regrettably, added to the individual’s SSR, and we have not yet been able to identify a centralized source for such data. We have, however, verified with two field offices that the data are collected because they are material to benefit payments, due to deeming

<sup>286</sup> Impacts on DI allowances, caseloads and benefits for SSI applicants can also be estimated. DI allowances will include allowances for any SSI applicants who are denied SSI payments because they failed the means test.

<sup>287</sup> Question 15 of the long form asks for the name, relationship, sex, date of birth, and disability status of everyone with whom the applicant lives. It also asks if anyone living with the applicant who is not married and under age 18 or between ages 18 and 21, not married, and a student receives any income, and, if so, the source, type, and monthly amount of the income.

of income for the support of family members.<sup>288</sup> Presuming the data are accessible, SSA would need to create an extract, and merge it to an SSR extract via SSN.

The evaluator could conduct two DID analyses with the three parenthood status groups: unmarried parents vs. non-parents, and married parents vs. non-parents.<sup>289</sup> Because most parents in AFDC/TANF families are unmarried, the impacts of TANF on allowances to unmarried parents are likely to be greater than on those to married parents. Thus, we would expect to find larger effects for the first comparison than for the second.

As with the DID analysis based on TANF status at application, this analysis should be performed by sex and narrow age category at the state level, with aggregation of findings to broader age categories and the national level as required.

Sub-analysis could consider the age of a parent's youngest child, and whether any child is a SSI participant.<sup>290</sup> Other things constant, effects of TANF on SSI allowances are likely to be smaller for parents of very young children (those with children under two are exempt from work requirements by federal law), and for parents whose youngest child is nearing age 18. Effects on allowances for parents of children who receive SSI might be larger or smaller than for other parents. Some of these parents will be exempt from work requirements, which would reduce the incentive to apply. This might vary substantially by state. The opposite result may be found, however, because adult disability may be positively correlated with child disability among families targeted by TANF, and because parents of SSI recipients will likely be more familiar with the SSI program, application process, and resources available to help in obtaining benefits. The sub-analysis could use a base parent group (e.g., married parents with youngest child between the age of 6 and 12 and not participating in SSI) as the comparison group for the purpose of testing hypothesis concerning the relative impacts on the various parent groups.

### *Estimators Based on Parent Characteristics (Children)*

For children, there is substantial information about parents in the administrative records to distinguish among those most and least likely to be in AFDC/TANF target families. The two most useful pieces of information are likely to be whether the child lives with both parents ("parental status"), and parent Social Security earnings in the period just prior to application.<sup>291</sup> For allowances, caseloads and benefits, a four-way classification scheme might be worthwhile --

<sup>288</sup> Although SSI does not provide support for dependents, it does deem a limited amount of other family income as necessary for the support of other family members, so that it is not reduce the SSI payment amount.

<sup>289</sup> We have not divided the non-parent comparison group into married and unmarried non-parent groups on the assumption that sample sizes in the married non-parent group would be small.

<sup>290</sup> As discussed in a previous footnote, the ages of all relatives the applicant lives with are obtained upon allowance.

<sup>291</sup> The parent(s) of an SSI child applicant or recipient can be identified on the SSR by flagging all records in the SSR that have the same Housed Under Number (HUN) as the child applicant/recipient and then pulling those records where the Master File Type (MFT) is equal to XM (ineligible mother) or XF (ineligible father). For childhood disability cases, the HUN for the child and ineligible parent(s) is equal to the child's SSN. Parents include stepparents and adoptive parents. Earnings data would need to be obtained from the Master Earnings File. It is important to use earnings prior to application for classification purposes because the onset of child disability may reduce parental earnings. The evaluator could, instead, use family income deemed available for the child for classification purposes. Deemed income is available on the SSR, although just for allowed cases. Hence, estimates of TANF impacts on applications would need to be based on parental status only.

by parental status crossed with a categorical variable for “high” or “low” Social Security earnings as a percent of poverty income (e.g., above or below 150 percent of poverty) -- assuming that sample sizes will support the analysis.<sup>292</sup> Those with positive deemed incomes would be the comparison group for those with no deemed incomes within each parental status class.

### **3. DID Estimators for the Combined Impacts of TANF and DA&A Reforms**

To produce estimates of the combined impacts of TANF and DA&A reforms on program outcomes for adults, the evaluator will need a target group whose members are potentially affected by either one, or both, of the reforms, and a comparison group whose members are affected by neither. As discussed above, the addiction disorder indicator can be used to identify allowances most likely affected by the DA&A reforms. Thus, the evaluator could identify the target group allowances as those who were receiving AFDC/TANF payments at the time of application and/or were determined to have an addiction disorder.<sup>293</sup> All other cases would be in the comparison group. Alternatively, if feasible, the evaluator could include those who are parents and/or those with addiction disorders in the target group, and all others in the comparison group.

It would be useful to produce DID estimates for subgroups within the target group: those affected by TANF only, those affected by DA&A reforms only, and those affected by both. Three subgroup analyses will be of interest:

- Comparison of outcome changes for those affected by both reforms to changes for those affected by TANF only will provide an indication of the extent to which the DA&A reforms dampened the impact of TANF.
- Comparison of outcome changes for those affected by the DA&A reforms only to changes for those affected by both reforms will provide an indication of the extent to which the TANF reforms dampened the impact of the DA&A reforms.
- Comparison of outcome changes for those affected by the DA&A reforms only to those not affected by either reform will provide an estimate of the impact of the DA&A reforms on allowances to those not affected by TANF.

---

<sup>292</sup> We assume that the vast majority of child applicants from AFDC/TANF families would have no deemed income. This assumption could be assessed by analysis of the deemed incomes for those identified as receiving AFDC/TANF income at the time of application.

<sup>293</sup> For young adults, the evaluator might add any cases likely to have been formerly in the target group for the SSI child reforms.



#### **4. DID Estimators for the Combined Impacts of TANF and SSI Child Reforms**

The approach developed here parallels the approach developed for the analysis of the combined impacts of the TANF and DA&A reforms on program outcomes for adults. To produce estimates of the combined impacts of TANF and SSI child reforms on SSI outcomes for children, the evaluator will need a target group whose members are potentially affected by either one, or both, of the reforms, and a comparison group whose members are affected by neither. The evaluation could use the previously discussed scheme, developed by RAND (1998), to identify allowances to children in the groups most likely affected by the child reforms. Thus, the evaluator could identify the target group allowances as those who were receiving AFDC/TANF payments at the time of application and/or those most likely to be affected by the child reforms. All other cases would be in the comparison group. Alternatively, if feasible, the evaluator could use parent status rather than AFDC/TANF payment status at application.

It would be useful to produce DID estimates for subgroups within the target group: those affected by TANF only, those affected by the child reforms only, and those affected by both. Three subgroup analyses will be of interest:

- Comparison of outcome changes for those affected by both reforms to changes for those affected by TANF only will provide an indication of the extent to which the SSI child reforms dampened the impact of TANF.
- Comparison of outcome changes for those affected by the SSI child reforms only to changes for those affected by both reforms will provide an indication of the extent to which the TANF reforms dampened the impact of the SSI child reforms.
- Comparison of outcome changes for those affected by the SSI child reforms only to those not affected by either reform will provide an estimate of the impact of the SSI child reforms on allowances to those not affected by TANF.

#### **5. Analysis of Non-Citizen Reforms**

PRWORA provisions would have ended SSI eligibility for non-citizens, with some exceptions, as of August 1997. The Balanced Budget Act restored eligibility for all legal aliens, provided they meet other program criteria, with one major exception. Most individuals who become legal immigrants after August 22, 1996 will not be eligible for SSI even if they otherwise qualify until they have accumulated at least 40 qualifying quarters of work.<sup>294</sup> Because all of these provisions apply to those who would otherwise qualify on the basis of either age or disability, the impact evaluation should consider adults of all ages. There may also have been an impact on children,

---

<sup>294</sup> There are many special categories of immigrants for which there are special exceptions. See “Impact of Welfare Reform Changes on Qualified Aliens,” *Social Security Talking Points*, September 1997.

but in 1995 there were only about seven thousand child SSI recipients who were classified as non-citizens.<sup>295</sup>

While the PRWORA provisions that were reversed in the BBA were never implemented, there may well have been a short-term impact on applications from, and allowances to, non-citizens. There was also likely a substantial impact on naturalizations, so any declines in non-citizen applications could be due, in part, to the increase in naturalizations. It is also possible that some non-citizen recipients were induced to leave the rolls, in anticipation of losing benefits, or due to misunderstanding about their eligibility. Further, even post-BBA there may be fewer applications from pre-PRWORA immigrants because of misunderstandings about eligibility, or perhaps loss of immigrant trust in SSI.

The SSR identifies applicants who are immigrants and provides information about their immigrant status (including whether they have become naturalized citizens), date of immigration, and country of origin. This immigration status field was updated following the 1996 reforms reflect naturalizations that would have been material to continuing eligibility had the initial reform not been changed.<sup>296</sup> It appears that we cannot distinguish between those who became naturalized citizens before application and those whose status was changed to “naturalized” after the reforms.<sup>297</sup>

Because date of immigration is material to eligibility under the BBA, an evaluation would need to follow immigrant applications and allowances by date of immigration to detect any effect beyond effects experienced in the period between PRWORA and the BBA. A complicating factor in the analysis is that immigration flows have varied substantially in recent years, especially due to the large numbers who received legal immigrant status under the Immigration Reform and Control Act of 1986 (IRCA), from 1988 through 1994. Data available from the Immigration and Naturalization Service (INS) can be used to control for immigration by fiscal year. Compiled data for each year include legal immigration by age (under 15, 15 – 29, 30 – 44, 45 – 64, and 65+) and sex, and by country of origin. Individual data can be obtained on data tapes, from as early as 1972.<sup>298</sup> The latest year for which data are now available is 1996. Tabulated data are available from 1992 forward on the INS website, and have been published in the *Immigration Yearbook* since 1984.

While much could be done with the immigration data, it would be worthwhile to perform some fairly simple analyses first. We recommend constructing cumulative adult applications and allowances per capita for recent immigrant cohorts by years since immigration. We believe this

<sup>295</sup> Pickett and Scott (1996, Table C). The corresponding figures for disabled adults is 543,600 and for aged adults is 236,600.

<sup>296</sup> The Alien-Refugee Indicator (ALIEN) on the SSR categorizes provides codes that can be used to classify all applicants as U.S. born citizens, naturalized citizens, and non-citizen immigrants in a variety of categories. The naturalized citizen category includes children of citizens who were born outside of the country, but presumably this is a small share of all naturalized citizens. ALIEN-RD provides the date of immigration, and ALIEN-CNTRY provides the country of origin.

<sup>297</sup> We are continuing to check on this. This information would allow the evaluator to estimate the impact of the reforms on citizenship among those who were already SSI recipients.

<sup>298</sup> The information on data availability was provided by Eloise Thornton of the INS Statistics Division.

could be done through the use of data from a single 10% SSR File from a post-reform year.<sup>299</sup> Based on an analysis of 1995 applications in the 1% SSR Longitudinal File, the 10% File should contain records for about 5,000 non-citizen applicants in the aged category in 1995, and 11,800 in the blind or disabled adult category, but only about 300 in the blind and disabled child category.<sup>300</sup> We expect the number of cases from each recent immigrant cohort to be substantial because of the rapid growth in non-citizen recipients that led up to the reform legislation. This is verified by an analysis of the December 1992 10% SSR File that showed that 16.5 percent of immigrant SSI recipients at that time had applied for SSI within 12 months of immigration, and 53.5 percent had applied within five years (*Scott, 1993*). Hence, we expect a majority of the applicants from each recent year to have entered within the last five years. If half of the immigrant applicants in 1995 entered within the last five years, then the 10% SSR would include an average of over 1,000 adult blind or disabled applicants and about 500 aged applicants from each of the previous five annual immigrant cohorts.

Application and allowance counts for each cell would be obtained from the SSR, and the denominator for each cell would be obtained from INS data. A table shell for the total results appears in **Exhibit 6.4**, which could be replicated for age-sex and country of origin subgroups. This would establish per-capita cumulative application and allowance schedules by cohort for the pre-reform cohorts. We would expect a downward shift in the cumulative application and allowance schedule for the 1996 cohort, and near elimination of applications and allowances for many years to the 1997 cohort. The application and allowance experience of earlier cohorts can also be used to project future effects on SSI outcomes under the assumption that all those who immigrate after August 22, 1996 remain ineligible for at least nine years. We might also see a flattening of the schedule for earlier cohorts in 1996 and 1997, possibly followed by a return to growth in 1998.<sup>301</sup>

It would be problematic to disentangle the effects of the non-citizen policy changes from the other reforms, because it is difficult to define comparison groups for immigrant applicants. The main reason is a factor that affects applications from immigrants but not native-born citizens: the flow of immigrants into the country. The evaluator could use INS and Census information on immigrant and non-immigrant populations to construct series for applications per capita for both groups, but this is at best a partial solution because date of immigration is key to the impacts. We do not have a recommendation for a comparison group. Given the limited nature of the non-citizen reforms under the BBA, more effort to address this issues does not seem warranted.

The post-PRWORA program status of immigrants who were SSI recipients just prior to the passage of PRWORA is also of interest. As a simple first step in assessing whether PRWORA resulted in a temporary or permanent reduction in participation by this group, the evaluator could tabulate monthly eligibility for SSI recipients who were eligible for benefits in December 1995,

<sup>299</sup> Scott (1993) used the 10% SSR for December 1992 to produce tables for months from immigration to SSI application for SSI recipients who were on the rolls in that month.

<sup>300</sup> Pickett and Scott (1996, Tables J, K, and L).

<sup>301</sup> If citizenship status at time of application can be distinguished from citizenship status after passage of PRWORA, a second set of exhibits could be developed to evaluate the impact of PRWORA on naturalizations among SSI applicants and recipients. Because there is a minimum five-year waiting period before naturalization, this analysis would only consider cohorts entering at least five years before the end of the analysis period.

from December 1995 forward, for three groups: immigrants who were not naturalized citizens at application, immigrants who were naturalized at application, and native-born citizens. The two comparison groups could be selected to match the target group on several important characteristics, such as age, sex, year of first allowance, state of residence and impairment.

Interaction between non-citizen reforms and TANF or other SSA reforms is probably not worth assessing because of the relatively few immigrants who are affected by the change to the current policy. Statistics on the caseload for the DA&A reforms also show that interactions between the DA&A reforms and the non-citizen reforms must be small.<sup>302</sup>

---

<sup>302</sup> Of the 167 thousand SSI recipients who were designated as DA&A in March 1996 (the month before elimination of allowances to applicants whose DA&A is material to their disability) only 1.0 percent were known to be non-citizens. Of these, 40 percent were still (medically) eligible as of December 1997. See Lewin (1998a).

**Exhibit 6.4**  
**Percent of Immigrants Applying for SSI and Receiving Allowances, by Year of Immigration and Years Since Immigration\***

Years After Date of Legal Immigration	Year of Legal Immigration							
	1990	1991	1992	1993	1994	1995	1996	1997
<b>Percent Applying</b>								
Less than 1 year								
Less than 2 years								
Less than 3 years								
Less than 4 years								
Less than 5 years								
Less than 6 years								
Less than 7 years								
Less than 8 years								
Less than 9 years								
<b>Percent Receiving an Allowance</b>								
Less than 1 year								
Less than 2 years								
Less than 3 years								
Less than 4 years								
Less than 5 years								
Less than 6 years								
Less than 7 years								
Less than 8 years								
Less than 9 years								

\*Shaded cells correspond to 1996.

## C. Pooled Time-Series Analysis of DID Estimates

### 1. Objectives

One significant limitation of the DID analysis described above is the assumption that “other factors” affect target and comparison group outcomes proportionately (e.g., a percentage point increase in the unemployment rate increases applications from both groups by the same percent). This assumption may be wrong. Pooled time-series analysis (PTS analysis) could be used to adjust state-level DID estimates for changes in other factors that can be observed at the state level. PTS analysis can also be used to assess the relationships between specific state policy changes and the DID estimates.

As mentioned previously, PTS analysis of applications and allowances alone would likely produce unsatisfactory information about the impact of TANF, because it is too difficult to adequately model the effects of the various state-level factors, including the policy changes. The

analysis proposed here would use state-level DID estimates as the first-cut estimate of the impacts of the TANF reforms, as well as of the combined reforms, and use PTS analysis to better understand variation in the estimates across states and over time.

Some of the DID estimators described above can be used for applications, but others cannot. The discussion below refers to allowances, but the approach can equally be applied to applications when DID application estimates are available. The approach could also be applied to caseload and benefit estimates, but findings are likely to be stronger for the allowances.

PTS analysis adds substantially to the data requirements for the DID analysis alone. It will require construction of historical DID series for a number of years before the reforms of interest. It would also be preferable to use quarterly, rather than annual series. It will be necessary to collect state-level explanatory variable data – building on the data that we have collected for this project through 1996.

## **2. Technical Specification**

### *Dependent Variables*

The analysis would use state-level DID estimates as dependent variables in pooled time-series models. The historical series would be constructed just like the post-reform series, using the last pre-reform period as the base. For allowances, the series will be the change in allowances from the base period to each specified period that is not explained, in a proximate sense, by concurrent changes in comparison group allowances. The value for the base period will be zero by definition. We suggest adding base period target group allowances in each state to the every value in a state's series. The resulting “normalized” series for each state will estimate target group allowances in each year after holding constant all factors controlled for by the DID analysis at base year values. The one-period change in the logarithm of the normalized series might then be used as the dependent variable in the analysis.

The aggregate DID series for each state will control for the effects of population growth and aging, as well as other factors, on target group allowances. Nonetheless, it would be useful to estimate separate PTS models for young women, young men, older women and older men, because the influence of other factors on the DID series will likely vary by age and sex. A single child model might be sufficient, but analysis by broad age group and/or sex might be warranted. The DID estimates for narrower age groups should be reviewed to determine reasonable age breaks.

### *Explanatory Variables*

We distinguish between two types of variables: control variables, to capture factors of no direct interest to the evaluation that may have different impacts on target and comparison group allowances; and program parameters, which capture state-level policy changes of interest to the evaluation. For the former, we are more concerned about controlling for the other factors than we

are about using a parsimonious specification with easily interpreted coefficients. We recommend including control variables from each of the following groups:<sup>303</sup>

- *Labor market* – we suggest continuing to experiment with the three measures used in the previous chapter: the unemployment rate, trade employment per capita, and the labor force participation rate.
- *General assistance* – we suggest using dummy variables to “dummy out” three years of data for a state when the state makes a major change to its GA program.<sup>304</sup> An alternative approach we have used in the past is to include an crude estimate of the impact of the GA policy change on per capita changes in the GA caseload.<sup>305</sup>
- *Other* – while we have attempted to use a wide variety of other control variables in our past modeling efforts for DI, SSI and AFDC, none have consistently proven to have explanatory power (see *Lewin, 1998b*). Perhaps the most promising of the others are those that have served as crude proxies for single-parent families -- vital statistics on marriage, divorce and out-of-wedlock births. While it is clear that changes in family structure are an important determinant of AFDC/TANF caseloads, and there is some evidence that they affect SSI caseloads as well (*Lewin, 1995a and b*), state-level measures of these variables are poor except in Census years. Similarly, state poverty rates would likely have explanatory power were it not for the fact that they are estimated from survey data. We have also attempted to use a variety of Medicaid program provisions in past modeling efforts, but have not obtained significant results. Medicaid expansions for women and children seem especially relevant, and have been shown to be significant in analysis of survey data on AFDC participation, but our efforts to use measures of these expansions in PTS analysis of AFDC caseload data did not yield statistically significant results. In the AFDC models we have also tried variables to capture possible implementation of laws concerning child support enforcement, paternity establishment, and abortion, but have obtained no significant effects.

There are two sets of AFDC/TANF program parameters to consider. One set consists of three parameters that we have used in the past to characterize the budget constraint for an AFDC family of three (*Lewin, 1997*): the maximum monthly benefit for a family of three, the average tax and benefit reduction rate, and a measure related to the program’s gross income limit.<sup>306</sup> There was, unfortunately, a temporary interruption in the collection of the data used to construct these variables in 1997.<sup>307</sup> As documented in *Appendix Exhibit D.19*, we have also constructed a

<sup>303</sup> Details of the relevant series that we have constructed for this and earlier efforts appear in Appendix Exhibit D.19.

<sup>304</sup> We applied this approach in the previous chapter, to four years, but did not find substantial findings after the first three.

<sup>305</sup> This variable was used in the hazard analysis reported in the previous chapter, and is documented in *Appendix Exhibit D.19*.

<sup>306</sup> We have already developed these measures through 1996, taking into account the interactions of AFDC, Food Stamps, and the Earned Income Tax Credit.

<sup>307</sup> The Congressional Research Service (CRS) has historically produced semi-annual reports summarizing the AFDC financial eligibility and benefit computation rules used by each state. However in 1997, the CRS did not survey states about these rules, because most states were in the process of implementing new eligibility and benefit rules as part of their TANF programs. Consequently, only limited information is readily available about

series of dummy variables to represent miscellaneous AFDC provisions that states have implemented in the recent past, including certain work requirements and limits to payments for children born while the mother is an AFDC recipient.

The second set of program parameters represents the state's implementation of TANF features. Developing such a set will be a major challenge, but is facilitated by the considerable efforts under way to document TANF reforms.<sup>308</sup> A simple scheme would be to develop dummy variables for implementation of work requirements and the five-year lifetime limit. It would also be valuable to represent the severity of sanctions. We found, for instance, that some states are planning to limit sanctions to elimination of benefits for the family's adults. Dummy variables to represent exemptions for people with disabilities from work requirements and the five-year limit might also be constructed. A recently completed study documents current state exemptions for TANF recipients with disabilities (Urban Institute, 1998). Changes to these features are likely to occur in some states in the future, but we do not know if such changes will be documented. SSA might find it advantageous to support efforts to do so. Information on new efforts to divert TANF recipients into SSI could also be documented in this way. In addition to these variables, it will likely be necessary to include dummy variables for miscellaneous reforms implemented in a very small number of states.

We are somewhat pessimistic about identifying the impacts of specific TANF reforms through the inclusion of these program parameters. This stems in part from the fact that we did not find demonstrable effects of changes in AFDC program parameters on SSI applications in our own PTS analysis of annual pooled applications for 1988 to 1996.<sup>309</sup> It may be, however, that the effects were obscured because variation in these variables was limited during the period and because major, imperfectly measured, changes in other factors (e.g., policy changes) dominated changes in the application series over this period. We would expect the DID estimates to be much more sensitive to AFDC/TANF program parameters. Use of quarterly data might also substantially improve the ability to detect the effects of changes in these parameters.

An alternative to use of the program parameters that capture the reforms is to classify states into three to five groups, based on a careful qualitative assessment of the reforms and their likely impact on SSI recipients. For instance, the group of "high impact" states might include those that: don't exempt TANF recipients with disabilities from work requirements and time limits; have shorter time limits than required by PRWORA; sanction the entire TANF unit, not just the parent(s), for non-compliance; and have diversion programs that require and actively support application to SSI for those with disabilities. Low impact states might be those that have none of these features, and all other states might be classified in one to three intermediate groups. This approach recognizes that it will be very difficult to disentangle the effects of specific reforms, but might be the best approach to identifying the joint effects of all reforms.

---

the financial eligibility and benefit computation rules used by states in 1997. The CRS resumed its survey of states in 1998, producing two reports, one in June and the second in October, on the eligibility and benefit rules used under state TANF programs.

<sup>308</sup> See Section VII for further discussion.

<sup>309</sup> The models we report in the previous chapter do not include these parameters, but this is because we found no credible evidence of effects in some initial models. The reported models do include dummy variables for a few particularly notable AFDC waivers, but again we found very little evidence of an impact.



The PTS analysis will implicitly include state fixed effects because the DID estimates are changes from the base period to the current period within each state.<sup>310</sup> That is, it would rely on cross-state co-variation between changes in explanatory variables and the percent changes in target group allowances that are captured in the DID series.

### **3. Interpretation**

Once estimated, the models could be used to produce:

- Adjusted estimates of the impacts of TANF on child and adult allowances, nationally and for each state;
- Adjusted estimates of the impacts of TANF and DA&A reforms on adult allowances, nationally and for each state; and
- Adjusted estimates of the impacts of TANF and SSI child reforms on child allowances, nationally and for each state.

Adjusted state estimates would be obtained by using the control variables and their coefficients to remove the variation in each DID series that is explained by the control variables. Any estimated effects of the program parameter variables would be retained in the adjusted estimates. The state estimates could be added to obtain national estimates.

The coefficients of the program parameter variables would provide some evidence of how the various features of TANF reforms affect SSI allowances. We would expect the provisions represented in these parameters to capture only a fraction of the effects of TANF, because it is simply not possible to fully capture the richness of cross-state variation in TANF programs, and in their target populations, that is relevant to the impacts of TANF on SSI outcomes. Nonetheless, the analysis may show that certain TANF features are particularly important in determining SSI outcomes.

## **IV. ANALYSIS OF MATCHED CENSUS/SSA DATA**

### **A. Overview**

SSA has linked data from the 1984, 1990, 1991, 1992 and 1993 Surveys of Income and Program Participation (SIPP) and the 1991 and 1994 Current Population Surveys (CPS) to SSA administrative data.<sup>311</sup> Future matches of both surveys are anticipated. This option would use these data to:

---

<sup>310</sup> The DID estimates are changes even before we convert to changes in the logarithms of normalized series (from the base period to the current period). The latter transformation converts everything to approximate annual percentage changes, making the series comparable across states of all sizes.

<sup>311</sup> Section F-1 of the RFTOP specifies that the SIPP files have been linked to the Supplemental Security Record (SSR), the Master Beneficiary Record (MBR), and the Summary Earnings Record (SER) and that the CPS files have been linked only to the MBR and SER. We assume, however, that survey files already linked to any of

- Estimate the impact of TANF reforms on SSI applications, allowances, caseloads and benefits, given the SSA reforms, at the national level during the observed post-period. Estimates of impacts on DI outcomes for SSI applicants would also be produced (Section B); and
- Estimate the combined impacts of all reforms on SSI caseloads and benefits in the post-period (Section C).

We used the 1990 – 1993 matched SIPP/SSA files to estimate the hazard models for SSI applications and allowances that are presented in the previous chapter. This analysis can be extended to study the impact of TANF on SSI outcomes. There is, however, an important *caveat*. It will be problematic to attribute estimated shifts in SSI applications from, and allowances to, those in the target population for TANF to TANF reforms themselves. As seen earlier, such shifts occurred before TANF. There are several explanations for these pre-TANF shifts, but our ability to discriminate among them is very limited. Future analysis is likely to encounter similar ambiguities.

Nonetheless, it would be useful for policymakers and planners to know when shifts from TANF to SSI are occurring, how large the shifts are, and the potential implications of the shifts for caseloads and costs. The estimates produced would also complement and validate the national estimates produced using the administrative data alone. A main advantage over the administrative data is the availability of extensive information on the characteristics of SIPP respondents, including family characteristics and past participation in AFDC – information that can be used to better define target and comparison groups.

The second part of this option addresses the need to evaluate the impacts of all reforms. We present a method that could use either the SIPP/SSA or CPS/SSA matched data. The approach would predict counterfactual caseloads in the post-period, using cross-sectional models estimated in the pre-period, and compare the size and characteristics of the actual and counterfactual caseloads. Actual and counterfactual benefits would also be compared. Again it will be problematic to attribute differences in the actual and counterfactual outcomes to the combined effects of the policy changes, exclusively. Differences in the characteristics of those in the actual and counterfactual caseloads should provide substantial information about how important the policy changes were. Cross-state analyses of differences between the actual and counterfactual caseloads may also be useful for this purpose.

## **B. The Impacts of TANF**

### **1. Objectives**

The primary objective of this analysis is to estimate the impact of TANF reforms on applications and allowances from 1996 through the end of the observation period, using multivariable

---

these records can potentially be linked to any other SSA administrative file. If this is incorrect, then what might be done with CPS data is more limited than we suggest in the text. The CPS data for 1991 and 1994 have been previously analyzed by Weaver (1997).

econometric models applied to the matched SIPP/SSA data. The estimates are likely to be confounded by the effects of other unknown factors that may shift participation from TANF to SSI during this period (or vice versa), but the methodology is expected to produce reasonably accurate estimates of changes in transitions from TANF to SSI for all reasons during the post-period.

A secondary purpose is to estimate the effects on SSI caseloads, SSI payments, DI beneficiaries, and DI benefits through the end of the observation period. This would be accomplished by estimating a series of auxiliary equations for allowed cases and using them to predict these outcomes for simulated counterfactual SSI allowances.

## **2. Data**

The econometric analysis would use pooled SIPP/SSA data. We recommend adding the 1996 SIPP panel to the 1990 – 1993 panels used in our analysis. The 1996 panel is the last panel to have its initial interview before PRWORA. The next scheduled SIPP panel will begin in 2000, and waiting for it would be of limited value because a considerable number of individuals who were at-risk for first SSI applications or allowances before PRWORA might have already become SSI applicants and recipients.<sup>312</sup>

In what follows, we assume that the 1990 – 1993 panels and the 1996 panel are available, and that SSI application and allowance data are available through the end of 2001.<sup>313</sup> By that date, five-year lifetime limits are likely to have been reached for a substantial number of TANF recipients. An earlier evaluation would be limited in its ability to capture the effects of lifetime limits, but may nonetheless find significant impacts because of work requirements, diversion efforts, and other TANF changes.

The sample for the analysis of the adults would be limited to SIPP respondents who: are at least 18 at the time they are first observed in the survey data; have low incomes; and are at-risk for SSI application or receipt. In our earlier analysis we used family income of 400 percent of the poverty standard as the income maximum, based on analysis of pre-application family incomes among recent SSI applicants. We used 40 as an upper age limit, in part because a large majority of adult AFDC recipients are under 40, but also because we wanted to be sure we had manageable sample sizes for the analysis. Our experience suggests that increasing the upper age limit to 50 would be feasible. It would also capture many more transitions to SSI, including transitions among parents who are over 40 when they are first observed. We recommend estimating separate models for women and men. Child samples would include those who are under 18 and still at-risk for SSI.

---

<sup>312</sup> The evaluator could also add the 1984 panel, but this may be problematic because of changes in SIPP data collection that occurred between 1984 and 1990. A possibly important advantage of adding the 1984 SIPP is that it would allow estimation of long-term duration effects in the period prior to the recent reforms, but interim SSA and non-SSA reforms may mitigate the value of that information.

<sup>313</sup> Availability of allowance data lags availability of application data due to the several months it usually takes to process applications (much longer in a few cases). Hence, complete allowance data are likely to be available for a shorter period than complete application data.

In defining “at-risk” for our application analysis, we excluded anyone who had previously applied, regardless of their current SSI status, because the matched data do not include information on later applications by individuals whose first application was denied or who had obtained benefits but left the program later. A large share of applications in each year come from applicant’s who have applied previously.<sup>314</sup> It would be desirable to add information about later applications to the merged file, so that the analysis could be expanded to incorporate effects on repeat applications.<sup>315</sup>

For our allowance analysis, we defined the at-risk population symmetrically – those who had never received an allowance before – and analyzed only first allowances. We understand that information on repeat allowances is available in the matched data, so this could be changed. The at-risk population would then be those who are not eligible for an SSI payment when observed in SIPP, including those who have been eligible at some point in the past, and the analysis would examine all future allowances, including repeat allowances. Past receipt of SSI would likely be an important addition to the explanatory variables.

SIPP weights should be used in each of the analyses described below to ensure that sample estimates are unbiased for corresponding values in the population. Use of weights will have little impact on the key parameters of interest if the effects they represent are reasonably constant across observations, as the specification assumes. Even if the assumption is correct, however, it will be necessary to use weights in any simulations performed with the estimated models to ensure that simulation estimates are unbiased estimates for population values.

### 3. *Econometric Model*

We first present a model that is a modified version of the hazard (duration) model presented in Chapter 5. We then discuss the potential use of simplified, linear versions, of this model that are computationally less demanding and perhaps easier to interpret, but problematic in other respects.

#### *Hazard (Duration) Model*

The following is a modification of the discrete time logit model for SSI applications and allowances that was presented in the previous chapter:

$$\text{Equation 6.1:} \quad \ln[R_{it}/(1-R_{it})] = a_{dp} + b'X_i + d_t \Pr(D_{it}^*/X_i)$$

where:

- $\ln[.]$  is the natural log operator;

<sup>314</sup> In 1993, 0.7 million of the 2.3 million SSI applications filed (30 percent) were first applications (Pickett and Scott, 1996, p. 37).

<sup>315</sup> As mentioned previously, some repeat applications are “noise,” reflecting administrative decisions or the passed appeal deadlines. The evaluator may want to only include the first repeat application after a specified period since the previous application (e.g., 12 months).

- $R_{id}$  represents the conditional probability that individual  $i$  applies for SSI benefits, or receives an allowance, in period  $d$  after he or she is first observed in SIPP;
- $\alpha_{dp}$  is the “duration effect” at duration  $d$  for respondents to SIPP panel  $p$ . This set of parameters allows for a different shift in the hazard at each duration for each panel, and each parameter can be thought of as a panel and duration-specific intercept;
- $X_i$  is a (column) vector of explanatory variables that do not vary with duration -- characteristics of the individual when first observed in SIPP;
- $b$  is a vector of coefficients for the  $X$ s;
- $D_{it}^*$  is a an unobserved dummy variable indicating that the individual is in the target populations for both TANF and SSI at time  $t$ .  $Pr(D_{it}^*/X_i)$  is the probability that the individual is in the target population conditioned on the information in  $X_i$ ; and
- $d_t$  is a year-specific coefficient for  $Pr(D_{it}^*/X_i)$ .

In thinking about this model, it is important to keep in mind that duration is being measured from the first point of observation in SIPP, and that the characteristics are based on that observation. Conditional on these characteristics, events such as changes in family status and onset of disability that occur at a later date are viewed as random.

There are two differences between this model and the model previously presented. The first is that we have allowed for different duration coefficients for each panel. We think this is important for the pre-reform period because the large increase in allowances during that period swept an increasing number of respondents out of the at-risk pool from 1990 to 1993. This probably accounts for significant coefficients for some of the panel dummies, especially for children. In our analysis, we imposed the assumption that panel effects do not depend on duration, but this may be incorrect. The specification presented here is more general. The more restrictive specification could be tested against this one, and adopted if the restrictions are not rejected.

It should be noted that the model has no year effects. Year effects would be of interest because they would show shifts in applications over time that cannot be explained by the variables in the models. In fact, however, year effects are imbedded in the duration-panel specific intercepts. The latter could be translated into an equivalent set of year-panel specific intercepts, because an observation’s panel and duration uniquely determine the year of the observation. Thus, shifts in the intercepts associated with panel and duration may, in reality, reflect unidentified shifts associated with year. We cannot add a set of year intercepts to the specification because they would be exactly collinear with the duration-panel specific intercepts. If we impose restrictions on the specification, however, we can identify year effects. For instance, if we specify that duration effects are all zero and just include panel parameters, we could put in year dummies as well – symmetric to the specification we have presented, which implicitly assumes zero year effects. Alternatively, we could specify that the duration-panel specific intercepts are the sum of a panel parameter and a duration parameter, and then include additive year effects for all but a base year. This might be more reasonable, but year effects produced in this way would be

subject to the criticism that they rely on unverifiable restrictions to distinguish them from duration effects.

The second important difference between this model and our earlier model is inclusion of the term  $\mathbf{d}_t'Pr(D_{it}^*/X_i)$ . This term allows for shifts in the hazard in proportion to the conditional probability that the individual is in the target populations for both AFDC/TANF and SSI at time  $t$ . This probability is unknown because disability is measured imperfectly in SIPP, and because change in family and disability status can occur between the survey date and  $t$ .

Suppose, for the moment, that  $Pr(D_{it}^*/X_i)$  is observed. Then we could include it as a time-varying variable in the equation. The difference between the value of  $\mathbf{d}_t$  for a post-reform year and the value for the last pre-reform year, which we assume to be 1995, would represent an estimate of the shift in the hazard for someone with a high probability of being in the target group. If TANF induces SSI applications or allowances, the effects would be captured through changes in the values of the  $\mathbf{d}_t$  after 1995. If TANF reforms were the only reason for the post-1995 values to change, then the impact of the reforms could be evaluated by using the model to simulate counterfactual applications and allowances after setting each post-1995  $\delta_t$  value equal to the 1995 value. Results from the pre-reform period suggest, however, that other sources of temporal shifts in the hazard for those targeted by the reforms are likely to be confounded with the effects of TANF. Nonetheless, it would be useful to know how large any future shifts of this sort are because they represent changes in applications and allowances from the TANF target population, for whatever the cause.

While we do not observe the probability that a SIPP respondent is in both target groups in a specific year, we do observe the conditioning variables,  $X_i$ . If we could somehow map these into probability estimates, we could substitute the probability estimates for the probability term in the equation.

One simple way to do this is to specify that  $Pr(D_{it}^*/X_i)$  is a linear function of a subset of the variables in  $X_i$ , say  $Z_i$ :<sup>316</sup>

$$\text{Equation 6.2:} \quad Pr(D_{it}^*/X_i) = \mathbf{g}_t'Z_i.$$

Substitute this expression into the right-hand side of the main equation to obtain:

$$\text{Equation 6.3:} \quad \ln[\mathbf{R}_{id}/(I-\mathbf{R}_{id})] = \mathbf{a}_{dp} + \mathbf{b}'X_i + \mathbf{d}_t'\mathbf{g}_t'Z_i = \mathbf{a}_{dp} + \mathbf{b}'X_i + \mathbf{g}^*Z_i$$

where  $\mathbf{g}^* = \mathbf{d}_t'\mathbf{g}_t$ . This suggests that we simply replace the probability term in the original equation with a term that has time-varying coefficients on a subset of the  $X$ s. The models reported in Chapter 5 in which we included a family dummy interacted with year dummies are a special case of this specification, with just one  $Z$  variable. It would be necessary to set  $\mathbf{g}^*$  equal to zero in some “base” year to avoid exact collinearity with the  $X$ s. The obvious base year would be the last full pre-reform year, 1995. With this normalization, the  $\mathbf{g}^*$  for years after 1995 would

<sup>316</sup> Duration since the SIPP observation also seems likely to have an impact on the probability of being the target populations for both programs. If we were to include duration shifters in this specification, however, they would be indistinguishable from the panel and duration specific intercepts in the model.

represent shifts in the hazard for target group cases relative to other cases after 1995. For the counterfactual of “no shifts,” we would simply drop the post-1995 terms from the estimated equation.

The  $X$ s for this analysis could be selected from the individual characteristics used in the preliminary analysis. In principle, many of the variables in  $X$  could be included in  $Z$ , but the small number of SSI transitions observed in each year make this impractical. It might be worthwhile to experiment with adding a second variable to the family. The leading candidate for an additional variable is a dummy variable for any disability or health condition interacted with the family dummy.

An attractive alternative is to first estimate a probability model for being a member of a targeted family (i.e., having children) *and* having a disability in the year in which the respondent is observed, then using the predicted value from the equation to estimate  $Pr(D_{it}^*/X_i)$  up to an unobserved factor of proportionality; the latter can be implicitly incorporated in  $\delta_t$ . This would be very similar to the model we estimated with terms for interactions between the estimated probability of AFDC participation and year. The assumption would be that the conditional probability in year  $t$  is proportional to the probability in the base year.

There are many ways that “having a disability” could be defined for purposes of identifying adults with disabilities who are in families. Because the prevalence of “severe disability,” using our SIPP-based definition, is high among adult SSI recipients (70 to 80 percent, depending on demographic group), this would be a reasonable definition to use for adults. Child disability is poorly measured in SIPP, and the only practical definition for this purpose may be “any disability” for child cases.<sup>317</sup>

“Targeted families” (i.e., families targeted by TANF reforms) could also be defined in many ways. They should include AFDC families, but other low-income families (e.g., below 200 percent of poverty and mother-only families) should also be included. The latter group is important because TANF reforms might divert them from ever entering TANF.

Explanatory variables for adults might include a dummy for any disability along with the variables we used in our AFDC probability model: age, race/ethnicity, marital status, education, children in family, other adult in family, and age of youngest child.<sup>318</sup> Explanatory variables for children would not include a disability dummy, but would include characteristics for both the child (age, sex) and parents (presence of each in the family, race/ethnicity, education, and disability), again following the specification we developed for the AFDC probability model.<sup>319</sup>

We did not include variables associated with the respondent’s state in this specification, apart from the state effects. Our experience with time-varying state variables in our earlier analysis was not promising, although improvements in the specification might produce more useful findings. The state-level explanatory variables that would be candidates for inclusion are the same as those that would be used in the pooled time series analysis of administrative data (see

<sup>317</sup> We found that only 56 percent of child SSI recipients had a disability by this measure.

<sup>318</sup> See *Appendix Exhibit E.11*.

<sup>319</sup> See *Appendix Exhibit E.12*.

previous section). If the variables are developed for that purpose, experimentation with including them in the hazard analysis would require little additional effort.

We understand that staff at ORES have been developing a methodology for predicting SSI eligibility in the general population, stemming from the allowance models developed by Hu et al. (1997) using matched SIPP/SSA data. The prediction methodology might eventually be incorporated in analysis of welfare reforms. At a minimum it would be interesting to observe change in the number of AFDC/TANF recipients who are predicted to be eligible for SSI from the pre-reform SIPPs to post reform SIPPs. Predicted probabilities of SSI eligibility could also be used to replace  $Pr(D_{it}^*/X_i)$  in the hazard

### *Linear Probability Model*

The logistic model (Equation 6.1) has some desirable properties, but is computationally burdensome and its coefficients are difficult to interpret (see Chapter 5 for a discussion of the interpretation). An alternative is to replace this equation with a linear probability model.

The most straightforward change would be to replace the left-hand side of 6.1, the log-odds of the hazard rate, with the hazard rate itself,  $P_{id}$ . This equation would be interpreted as a linear hazard function, and could be estimated using the same sample as used for the logistic model. This sample includes an observation for each respondent in each period up through and including the period in which the respondent applies/receives an allowance, but not for later periods. Each coefficient can be straightforwardly interpreted as the effect of a unit change in the explanatory variable on the hazard.

Alternatively, one could estimate a linear cumulative probability function. This would be identical to the linear hazard function except that every respondent would have an observation for each sample period, including periods after the period in which the respondent applies/receives an allowance; the dependent variable is coded as unity in each such period. Under this specification,  $P_{id}$  is interpreted as the cumulative probability for respondent  $i$  at duration  $d$ .

Although computationally easy, the linear models described above have some shortcomings. First, the probabilities are not bounded between zero and one. Because SSI applications/allowances are relatively rare events, it might be that a substantial number of observations will have negative predicted probabilities. This may have little practical importance, however, because it is the coefficients, rather than the predicted probabilities, that are of interest. A potentially much more serious, related, problem is that the effect of a unit change in each explanatory variable on a probability in a linear model does not depend on the levels of other variables. In the logit model, the effect of the same change is proportional to the product  $P_{id} (1 - P_{id})$ , which is greatest when  $P_{id} = 0.5$  and approaches zero as  $P_{id}$  approaches either zero or one. The fact that probabilities are bounded between zero and one requires that the effect of a change in an explanatory variable diminish to zero as the probability approaches either extreme. The linear model is clearly misspecified in this regard, and might not fit the data as well as a logit model as a result, especially when there are many individuals with true probabilities near one of the extremes. Goodman (1977) demonstrates the linear and logistic



models can produce quite different results when the probability of an event occurring is less than 0.10 or greater than 0.90.

A final issue with linear probability models is that the disturbances are necessarily heteroskedastic. This well-known problem can be easily corrected through the use of weighted least squares.<sup>320</sup>

#### 4. Simulations

Once a model has been estimated, it can be used to simulate counterfactual applications from the first post-reform year (assumed to be 1996) forward. At the beginning of the first year, many of those “at-risk” for first SSI application or allowance will no longer be at risk, having already applied and or received an allowance. For those at risk, counterfactual applications and allowances can be generated with the model over the remaining years of the sample period (through 2001 under our assumption), by dropping the terms with time-varying coefficients (i.e., the time interactions). For each person still at risk, the evaluator would compute the fitted hazard rate ( $P_{it}$ ) for the first post-reform year and compare it to a random draw from a uniform (0, 1) distribution. If the randomly drawn number is below the fitted hazard, the case would be counted as a counterfactual application or allowance for that year. Such cases would then be dropped from the sample, and the process would be repeated for the next year. The simulation would iterate forward in this fashion until 2001.

The characteristics of individuals who “filed” counterfactual applications and or “received” counterfactual allowances can be compared to those for the actual applications and allowances. Perhaps more interestingly, the evaluator can compare the characteristics of the “marginal” actual applicants (those who would not have applied under the counterfactual) to the characteristics of the counterfactual applicants.<sup>321</sup> Presumably most of these cases will have been at high risk for being in the target groups for both TANF and SSI when observed.

The allowance model will yield analogous predictions of counterfactual first allowances. For these cases, it would be interesting to predict each of the following: SSI payment status for the remainder of the observation period, SSI payments, DI participation, and DI benefits (including dependent benefits) for the year in which the allowance is made as well as for later years. A simple way to do this is to match the counterfactual allowances to contemporaneous actual allowances of SIPP respondents on the basis of both characteristics observed in either SIPP or the administrative data. Matching variables might include year of allowance, year of birth, sex, marital status (when observed in SIPP), monthly DI benefit amount, highest level of education, race and ethnicity. Once a counterfactual case is matched to an actual case, the counterfactual case is assigned the SSI and DI eligibility and payment histories of the matched actual case from the date of allowance forward.

<sup>320</sup> See Greene (1990), Chapter 20.3.

<sup>321</sup> If  $s_a$  is a statistic (mean or percent) for a characteristic of those who actually filed an application,  $s_c$  is the corresponding statistics for counterfactual applicants, and  $k$  is the number of counterfactual applicants relative to the number of actual applicants (presumably less than one), then the statistic for the marginal group is:  $s_m = (s_a - k s_c)/(1-k)$ .

Because the sample of allowed cases observed in SIPP in each year is small, the matching variables will need to be prioritized, and the matching may not be very close. An alternative is to use characteristics observed in administrative data only, and match cases to a larger sample of allowed cases. A second alternative is to estimate simple econometric models for the outcome variables using data for actual allowed cases and explanatory variables such as those described above. The models might include:

- A logit model for monthly SSI payment status in each month after the allowance. Every month after the first allowance month for each individual until the last month of the sample period would be an observation. The model should include number of months since allowance along with the other explanatory variables. Some may leave SSI within a few months because they qualify for DI, while others may leave SSI for other reasons, and some may leave and return;<sup>322</sup>
- A multiple regression model for mean monthly SSI payments during eligible months;
- A logit model for concurrent DI allowances;
- For those who receive an DI allowance, a linear probability model for monthly DI eligibility in each month after the SSI allowance, analogous to the SSI payment status model;<sup>323</sup> and
- For those who receive an DI allowance, a multiple regression model for the initial monthly payment amount. Payments amounts for later months can of eligibility can be projected from the initial payment using SSA's cost of living adjustment methodology.

The estimated models could be used to predict monthly SSI payment eligibility, monthly SSI payments, DI allowance, and DI monthly payment amounts for the counterfactual allowances from the date of allowance through the end of the observation period.

More elaborate models for these variables could be developed from administrative data for large numbers of cases, but we think that additional effort in this area would have a relatively small payoff. Continuing eligibility models using administrative data could build on termination models that have been previously developed, but would not have several of the explanatory variables that are available in SIPP.

## **C. The Combined Impacts of All Reforms on Caseloads and Benefits**

### **1. Objectives**

The objective of this analysis is to estimate the impacts of all reforms, combined, on program caseloads and benefits, using the SIPP/SSA matched data. It is likely that this objective will not be fully attained because of the difficulty of controlling for various confounding factors.

---

<sup>322</sup> An alternative would to estimate a liner or logit hazard model for SSI exit, but this would require estimation of a reentry model, too, because SSI recipients often have interrupted spells.

<sup>323</sup> Because of the five-month DI waiting period, DI eligibility may be delayed for up to five months.

Nonetheless, the analysis will provide useful information about the causes of caseload changes over the post-reform period.

This analysis can readily be applied to both SSI and DI, although it is anticipated that impacts for SSI will be much larger than for DI. Comparison of estimated impacts for SSI and DI may be helpful in interpreting the findings. The presentation here assumes that this analysis would be applied to both programs.

## **2. Data**

There are two candidate data sources for this analysis. One is the SIPP/SSA data for each panel, including the 1996 and later panels. The next panel is scheduled to begin in 2000, so this would limit the analysis to comparison of actual and counterfactual caseloads in the year 2000 and, possibly, later years. The second is CPS/SSA matched data. The CPS is conducted every year, and also has the appeal of larger sample sizes. The SIPP data are nonetheless appealing because of better health, disability, and program participation data.

For purposes of discussion, we assume that the 1996 SIPP panel will be used for the pre-period and the 2000 SIPP panel will be used for the post-period. We discuss changes to the analysis if the CPS/SSA data were to be used – some required by the limitations of the CPS, but others allowed by its strengths – at the end of this section.

We recommend estimating separate models for adults (age 18 to 64) and children. Although we would prefer to split the adult sample by sex and broad age group, the number of SIPP respondents who are SSI recipients may be insufficient to support such an analysis. Among adult respondents in the 1993 SIPP sample, 484 were SSI disability recipients in January of 1993 (*Appendix Exhibits E.1 – E.4*). As discussed further below, we are especially interested in the group of recipients who received their first allowance in the last five years. Only about 290 of the 484 January 1993 SSI recipients in the 1993 SIPP panel are in this category.

## **3. Methodology**

This analysis would use cross-section data from the pre-period to develop prediction models for contemporaneous SSA disability program outcomes – SSI and DI participation and benefits. Once the models are developed, they would be applied to cross-section data from one or more post-periods to predict counterfactual outcomes – the outcomes that we would expect based on the pre-period prediction models and the post-period characteristics of survey respondents. These outcomes can be compared to actual outcomes from the post-period. The characteristics of actual and counterfactual recipients can also be compared, to assess the importance of the various policy changes and other factors in explaining the difference between the actual and counterfactual program outcomes. The models can be re-estimated using the post-period data, and changes in *per capita* outcomes can be decomposed into changes in the relationship between the outcome and individual characteristics and changes in the mean characteristics of the population.

### *Econometric Model for Adults*

The model for adults would predict four different outcomes for each individual in the sample: participation and payments for each of the disability programs during the year. Participation would be defined as eligibility for payment in any month of the year, and payment would be based on the amount actually paid for those months in which the individual was eligible. A refinement would be to predict months of eligibility for those who are eligible, which would be needed to compute mean monthly caseloads.

The model might be structured as follows:

- A multinomial model to predict four participation categories: SSI-only, DI-only, concurrent, and none.<sup>324</sup>
- Multiple regression models to predict benefits for those in each participation category. There would be four equations: an SSI equation would be estimated for SSI-only cases, a DI equation for DI-only cases, and an equation for each program for concurrent cases.

### *Explanatory Variables*

Differences between the actual and counterfactual outcomes will reflect the effects of factors that have not been controlled for by the explanatory variables in the model. Selection of these explanatory variables is, therefore, critical to the findings and their interpretation. An important challenge to this approach is that many variables that program participation in a cross-section are also influenced by changes in the economic and policy environment; i.e., they are endogenous.

The following variable types are candidates for inclusion in the participation equations. Those at the beginning of the list are clearly exogenous – not affected by factors that determine program outcomes. Those at the end are, in our judgment, the most likely to be influenced by the reforms.

- Age, sex, race and ethnicity. These basic demographic variables are clearly important determinants of program participation, and clearly exogenous;
- A set of dummy variables for SSI or DI participation more than five years earlier. This is a critical variable because when the next SIPP is started, in 2000, we will be five years into the post-reform period. We would expect the probability of participation given that the respondent had participated more than five years earlier to be smaller in 2000 than in 1996, because of the SSA reforms. The combined reforms may increase or decrease the probability of participation given that the respondent was not a participant more than five years earlier, because the TANF and SSA reforms have opposing effects on allowances. These dummies are “predetermined;” they are exogenous with respect to events of the last five years that may influence participation, but are not exogenous to earlier events.

---

<sup>324</sup> Our understanding is that it is now feasible to estimate a variety of multinomial models with four (or even more) outcomes that computationally more challenging than the formerly popular multinomial logit model, but that, unlike the later, do not have the “independence of irrelevant alternatives” problem. These models include multinomial probits. LIMDEP 4.0 offers several alternatives ([www.limdep.com](http://www.limdep.com)).

- Educational attainment. Because schooling levels among SSI recipients are low relative to those in the non-SSI population, educational attainment is likely to be a strong predictor of program outcomes (see Chapter 5). Endogeneity will not be an issue except for children and young adults.
- Disability and health status. These variables will clearly be strong predictors of SSI and DI receipt. Past research has shown, however, that self-reported disability and health measures are influenced by the policy and economic environment.<sup>325</sup> This may be a more serious issue for “any disability” measures than it is for “severe disability” measures.
- Place of residence. This could include state dummy variables (state fixed effects), and/or dummy variables for urban or rural residence. It is possible that some reforms, especially TANF, may influence people to change their place of residence – most likely to move across state boundaries.
- Marital and family status. It seems likely that there will be a relationship between marital and family status variables and SSI participation in a cross-section, holding other factors constant. These variables may, however, be influenced by the reforms. Indeed, among other things, the architects of TANF sought to reduce out-of-wedlock births and encourage the formation of two-parent families.
- Employment, earnings, and other non-program income. Current values of these variables will likely be predictive of current SSI participation, and it would be very desirable to control for changes in these variables because they are influenced by the state of the economy. They are, however, endogenous, both because changes in participation induced by the reforms are likely to be accompanied by income and employment changes in many cases.

SSA might find it useful to estimate a series of models, starting with models that only include variables near the top of the list, and progressively adding others. The more inclusive models would provide more information about the correlates of caseload changes, but the interpretation of the correlates would become more problematic.

Adjusting for economic change may be especially problematic. We do not know how strong the economy will be in 2000. Even if its strength is approximately equal to its strength in 1996, the history of the economy between 1996 and 2000 will be quite different than the history of the economy between 1990 and 1996. We will return to this issue below.

#### **4. Use of the Models to Analyze the Impacts of Reforms**

A simple way to use the estimated models for analysis of the impacts of reforms is to predict program outcomes for 2000 using the SIPP panel for that year and the model estimated with the 1996 panel data. Predicted outcomes can be compared to actual outcomes to obtain an estimate of the changes that can be “explained,” in a proximate sense, by all factors that have not been captured in the explanatory variables. This will include the policy changes of interest, but may

---

<sup>325</sup> See Bound (1991), Waidean, et. al. (1995) and Kubik (1997).

include other factors as well – particularly the economy, and perhaps unrelated policy changes. Comparison of the characteristics of simulated counterfactual program participants to actual program participants will provide some evidence of the importance of the policy changes in explaining caseload changes. For instance, a finding that the counterfactual recipients contain a smaller share of young mothers than the actual recipients would suggest that the TANF reforms contribute to the difference between actual and counterfactual recipients. Some of the characteristics of interest will be included in the explanatory variables, but others will not (e.g., TANF participation).

The model could be estimated separately for both years, and the estimates could be used to produce a more formal analysis of the proximate causes of outcome changes from 1996 to 2000 will be helpful. Consider the following linear model for a specific outcome in year  $t$ :

**Equation 6.4:**  $Y_i = b_t'X_i + e_i$

where:

- $Y_i$  is the outcome variable for individual  $i$ . The most interesting outcomes will be dummy variables for program participation;
- $X_i$  is a (column) vector of explanatory variables for individual  $i$ ;
- $b_t$  is a (column) vector of coefficients for  $X_i$  in year  $t$ . They could be estimates from a linear probability model, or could be derived from a linear expansion of a non-linear model around the mean of the  $X$ s in year  $t$ .<sup>326</sup> In a participation model, each element of  $b_t$  would estimate the change in the probability of participation associated with a unit change in the corresponding explanatory variable in year  $t$ , evaluated at the sample mean of the  $X$ s for year  $t$ ; and
- $e_i$  is the residual (prediction error).

Given the estimated coefficients, the change in the mean value of the outcome from 1996 to 2000 can be decomposed as:

**Equation 6.5:**  $\bar{Y}_2 - \bar{Y}_1 = b_2' \bar{X}_2 - b_1' \bar{X}_1 = (b_2 - b_1)' \bar{X}_2 + b_1' (\bar{X}_2 - \bar{X}_1),$

where: over-bars indicate variable means, the subscript 2 indicates values for 2000, and the subscript 1 indicate values for 1996.

Equation 6.5 decomposes changes in the mean of the outcome variables to changes in the coefficients, weighted by post year means of the explanatory variables, and changes in the means of the explanatory variables themselves, weighted by the pre year coefficients. If the explanatory variables are exogenous to the policy changes, the effects of the policy changes are captured by the changes in the coefficients. The SSI participation analysis might show, for instance, that:

---

<sup>326</sup> An intercept is implicitly included as the coefficient of a “constant” in  $X$ .

- A decrease in the coefficient of the dummy for participation more than five years before the current period, reflecting the impacts of SSA reforms on existing participants; and
- An increase in the coefficient on a variable indicating that the applicant is a parent, interacted with a dummy variable for no participation more than five years before the current period, reflecting the impact of TANF.

## **5. Adjustments for Changes in the Economy**

One way to at least partially control for the differences in histories of the economies for the five years preceding 1996 and 2000 would be to:

1. Construct state-level variables for the strength of the economy in the latter five-year period relative to the former five-year period;
2. In the model for 2000, fix the state dummy coefficients (fixed effects) at the 1996 estimated values and add the constructed variables to the equation. This would limit the shifts in the state intercepts to be proportional to the measure of the relative strength of the economy.<sup>327</sup>

## **6. Use of Matched CPS/SSA Data**

Use of the matched CPS/SSA data would have two distinct advantages over the use of the SIPP/SSA data: larger sample size and annual observations (assuming that SSA matches the data every year). The CPS typically has three times as many respondents as the SIPP. The annual data are advantageous for two reasons. First, they allow construction of the estimates for each year. Second, pooled analysis using multiple years of data may significantly improve the evaluator's ability to control for the effects of the economy. The main disadvantage is that the health and disability data are quite limited. This would not be an issue for models that exclude health and disability variables, because of endogeneity. Poor data on participation in programs other than SSI and DI would also be an issue in the comparison of characteristics for those in the actual and counterfactual caseloads.

## **7. Strengths and Limitations**

The clearest advantage to use of the matched Census/SSA data relative to the use of the administrative data is that we can observe characteristics to define target and comparison groups for the non-SSA reforms in a more satisfactory way. We can also control for other characteristics that are predictive of SSI applications, such as observed disabilities as well as past employment histories.

---

<sup>327</sup> If fixed effects for all states are included in the 1996 model, then an intercept should be added to this version of the 2000 model, to capture outcomes shifts not explained by changes in the economy or other explanatory variables in the model.

Small sample sizes for program entry from the various groups of interest are the most serious limitation. This clearly impinges on our ability to separate the impacts of SSA and non-SSA reforms. Lack of information to identify cases targeted by DA&A and SSI child reforms is also a limitation. Controlling for changes to the economic and state policy environment (outside the non-SSA reforms of interest) may also limit the usefulness of this analysis.

The fact that TANF reforms vary substantially in nature and timing across states will also be problematic for this approach. The number of observations from each state will almost surely be too small to draw any firm conclusions about impacts for that state. Hence, we may well miss large impacts in some states because of small average impacts over all. Inclusion of explanatory variables that characterize TANF reforms could help identify the reforms' impacts, but again we are concerned that idiosyncratic variation in behavior will hide even substantial effects.

#### **D. Validating and Improving the DID Analysis of Administrative Data**

Earlier in this chapter we developed a difference-in-differences (DID) methodology for comparing growth in allowances for applicants who reported receipt of AFDC/TANF at the time of application to growth in allowances for other applicants, holding age, sex and possibly other factors constant. We would prefer to classify applicants by whether they had ever received AFDC/TANF.

The matched SIPP/SSA data offer an opportunity to study the relationship between the report of AFDC/TANF receipt at the time of application in the administrative data and self-reported past receipt of AFDC/TANF.<sup>328</sup> The evaluator could use the matched data for SSI applicants to develop a model that predicts past receipt of AFDC/TANF, as reported in SIPP, from information that is observed in the administrative data, including receipt of AFDC/TANF at the time of application. Other variables to include would be age, sex, race/ethnicity, state, and time between SIPP observation and SSI application. Note that the SIPP observation may be before or after SSI application. At a minimum, this analysis would provide SSA with an indication of the extent to which the administrative data on AFDC/TANF receipt at application capture any past AFDC/TANF receipt.<sup>329</sup>

The estimated relationship could also be applied to the administrative data for the purpose of developing alternative target and comparison groups. The idea is to produce series that better approximate allowances to the target group “applicants who are former AFDC/TANF recipients” and to the comparison group “applicants who are not former AFDC/TANF recipients.” More specifically, the equation estimated with the matched data could be applied to predict the probability of past AFDC/TANF participation for each observation in the administrative data sample (including the many not observed in SIPP). Summing the predicted probabilities over all allowances made in a given year and state yields an estimate of the number of allowances made

<sup>328</sup> It is our understanding that the administrative report of AFDC/TANF receipt is not currently in the matched file, but it could presumably be added if the analysis described were believed to be sufficiently important.

<sup>329</sup> Earlier in this chapter we mentioned that the *SSI Annual Report* includes an estimate of the number of current SSI recipients who are former AFDC recipients. The number reported might be a significant understatement because it leaves out those who are former AFDC recipients but who were not receiving AFDC at the time of SSI application. The analysis described here would provide a clear indication of the magnitude of the bias.



to former AFDC/TANF recipients in the state in that period. Subtracting from total allowances in the same state-period yields an estimate of the number of allowances made to individuals who are not former AFDC/TANF recipients. The analysis could be done by age, sex and possibly other factors.

A significant limitation of this approach is that the relationship between past AFDC/TANF receipt and AFDC/TANF receipt at time of SSI application is likely to change as a result of the reforms. Eventually this could be checked with matched data from the 2000 SIPP. Another important limitation is potential misreporting of former AFDC/TANF participation in SIPP – in part because of the timing of data collection.

## **V. EVALUATING THE EFFECTS OF TANF USING EXISTING STATE WELFARE REFORM EVALUATIONS**

### **A. Overview**

A number of states implemented time limits and strict work requirements in conjunction with HHS waivers prior to the passage of PRWORA in August 1996. Several of these states have continued experimental evaluations of their programs and offer the best opportunity to assess the impact of these provisions on both adult and child family members. Experimental evaluations offer the unique opportunity to follow the paths of families randomly assigned to treatment and control groups. To the extent that their pattern of SSI applications are significantly different, it is reasonable to conclude that the difference is due to the program intervention.

SSA could work with these states and their evaluation contractors to identify the information that can be obtained under the existing design and to pursue the option of linking evaluation data with SSA administrative data. The latter would allow the contractor to follow SSI applications and allowances among treatment and control group members to supplement existing information as necessary.

### **B. Specific Opportunities**

As outlined in *Chapter 2* on Welfare Reform Evaluations, there are nine states that have experimental evaluations in place and offer the opportunity to track research group members' interaction with SSA programs. In *Exhibit 6.5* we identify the nine states for further consideration, their evaluation contractors, the program design, and the potential link to SSI. Five of these states: Connecticut, Florida, Indiana, Iowa, and Minnesota have also identified specific child impact measures for incorporation into their studies. Additional information on these state evaluations is available in *Appendix B*.

All of these states, except Minnesota, are implementing program treatments that include time limits and strict work requirements. Minnesota's approach tests the effect of strong financial incentives and time-triggered mandatory employment and training services. Each of the eight states employing time limits have adopted a different approach to the length of time benefits are offered, reasons for exemptions or extensions, as well as the mix of employment and support services offered to families. This natural variation is both an advantage and a disadvantage. On the positive side, it will allow SSA to explore the effects of a variety of approaches states can

take under TANF. On the other hand, SSA will need to exercise caution if it attempts to pool data across states. While the increased statistical power of a larger sample may be important to explore potentially small effects, the variation in specific state interventions is likely to make such pooling problematic.

### Exhibit 6.5 Evaluations of Non-SSA Reforms

Evaluation	Evaluator	Program Design	Administrative Sample Size	Administrative or Survey link to SSI Program Information
<b>Arizona</b>	Abt Associates Inc.	Time limits, family cap, extended transitional Medicaid and child care	5,829 welfare recipients	Analyses of survey data have been conducted on treatment and control group members who transitioned into SSI. SSNs are available for adults in administrative data.
<b>Connecticut</b>	Manpower Demonstration Research Corporation	21-month time limit, all earned income disregarded up to the poverty line, and modified family benefit cap.	6,090 welfare recipients	The survey included a question for whether the respondent or any household member received income from SSI, DI or aid for the disabled. An additional question was asked regarding if the income was for the respondent or someone else. SSNs are available for adults and children in administrative data.
<b>Florida</b>	Manpower Demonstration Research Corporation	Strict time limit (24 or 36 months out of any 60 months, depending on recipient characteristics and previous time on assistance) and generous income disregards in a low-benefit-level state.	5,430 welfare applicants and recipients	The survey included a question for whether the respondent or any household member received income from SSI, DI or aid for the disabled. An additional question was asked regarding if the income was for the respondent or someone else. SSNs are available for adults and children in administrative data, though children were not used in the evaluation
<b>Indiana</b>	Abt Associates, Inc.	24-month time limit and family benefit cap.	<ul style="list-style-type: none"> <li>• 10,706 in the recipient sample (includes individuals who were receiving assistance in May 1995)</li> <li>• 6,869 in the applicant sample (includes individuals who filed for assistance after May 1995)</li> </ul>	Survey included a question for whether the respondent received income from SSI, but no differential was made whether the SSI is for the child or adult. SSNs are available for adults and children in administrative data. Survey data is linked with administrative data.
<b>Iowa</b>	Mathematica Policy Research, Inc.	Strong work requirements (includes severe sanctions for non-participation) and expanded earnings disregards.	4,224 cases of in the Limited Benefit Plan. The number of cases in the Family Independence Program is not known.	There is a single income category for SSI and DI income for every individual in the household. SSNs are available for adults in administrative data.

**Exhibit 6.5 (Continued)**  
**Evaluations of Non-SSA Reforms**

<b>Evaluation</b>	<b>Evaluator</b>	<b>Program Design</b>	<b>Administrative Sample Size</b>	<b>Administrative or Survey link to SSI Program Information</b>
<b>Minnesota</b>	Manpower Demonstration Research Corporation	Generous income disregards, eligibility for supplemental benefits up to 140 percent of poverty, and intensive employment and training requirements for longer-term recipients.	14,369 welfare recipients	Survey included a question for whether the respondent or any household member received income from SSI, DI or aid for the disabled. An additional question was asked regarding if the income was for the respondent or someone else. SSNs are available for adults and children in administrative data.
<b>Nebraska</b>	Mathematica Policy Research	Intensive case management, time limits, extended transitional benefits	Approximately 7,200 TANF recipients	Linked administrative records on SSI receipt are available for all household members. Survey questions on SSI receipt are also asked. SSNs are available for adults and children in administrative data.
<b>Texas</b>	Texas Department of Human Services	Time limits, personal responsibility agreements	Between 15,000 and 20,000 TANF recipients	Small-scale surveys include questions regarding SSI receipt. SSNs are available for adults and children in administrative data.
<b>Vermont</b>	Manpower Demonstration Research Corporation	Strict work requirements, generous earnings disregards	10,997 welfare applicants	Survey included a question for whether the respondent or any household member received income from SSI, DI or aid for the disabled. An additional question was asked regarding if the income was for the respondent or someone else. SSNs are available for adults and children in administrative data, though SSA data would be needed to determine if the individual is a child or an adult.
<b>Wisconsin</b>	Institute for Research on Poverty	Diversion strategy, strict work requirements	Approximately 4,000 cases	No data is being gathered about SSI participation. SSNs are available for adults in administrative data.
<b>Employment Readiness Demonstration Project</b>	California State University-Bakersfield	Targeted services approaches for persons with multiple barriers to employment in eight counties in California	Approximately 1,500 Hard to Serve TANF cases	State administration data on SSI participation is available. SSNs are available for adults in administrative data.
<b>Welfare to Work Evaluation</b>	Mathematica Policy Research	Welfare to Work Strategies	Approximately 50,000 welfare applicants and recipients	Surveys will include questions regarding SSI income sources. Administrative data will be used, though the sites have not yet been selected.

**Exhibit 6.4** also identifies an experimental evaluation in California, the Employment Readiness Demonstration Project (ERDP), which offers the opportunity to explore the effects of mandatory work requirements and program services on individuals with multiple barriers to work. It may be interesting to explore the effects of a demonstration that focuses on individuals who may, as a group, have a greater probability of applying for SSI. All of the individuals in the research group are subject to a five-year time limit on receipt of cash assistance. Only those in the treatment group are receiving intensive services.

Finally, the newly funded Welfare to Work (WtW) Evaluation is still in its formative stages. HHS and its contractor, Mathematica Policy Research, have not yet selected sites to be included in the impact study. WtW will, by definition, focus services on the “harder to serve”. For this reason, SSA should consult with ASPE on the evaluation design and explore its relevance to questions of interest to SSA.

One way for SSA to pursue this option further would be to send each of the evaluation contractors a request for information. The request would specify SSA’s research questions and state SSA’s interest in funding add-on studies to welfare reform evaluations that would address those questions. Each contractor would be asked to describe its interest and capabilities for addressing the questions through an add-on to its existing evaluation(s). The response should include: 1) a technical description of what can be accomplished given access to the evaluation data and, if needed, matched SSA data; 2) a preliminary estimate of the level of effort required; and 3) a discussion of the willingness of the relevant state authority(ies) to permit use of the relevant evaluation data. The latter should clearly specify any technical or contractual conditions that the state authority(ies) would require. With this information in hand, SSA would be in a better position to decide which add-on evaluations to pursue further, and how to pursue them.

## **VI. STATE CASE STUDIES**

### **A. Overview**

SSA can supplement information it gathers through experimental studies by conducting case studies of specific states using both quantitative and qualitative methods. SSA can use state administrative data, SSA administrative data, and survey research data to track the SSI applications of current and former TANF recipients and the SSI allowances of former TANF recipients. This could be expanded to include difference in differences (DID) analyses that would validate the findings from the DID analyses described in our first evaluation option, based on administrative data alone. Qualitative case study data can be used to provide contextual information regarding the TANF program and policy initiatives that influence the movement of clients from TANF to SSI, the experience of SSA field offices and State DDSs, as well as the perceptions of state and local advocates and interest groups. SSA can implement this option by building on existing work in progress and by conducting its own tracking efforts and case studies.

## **B. Data Sources**

There are two primary options for building on existing work. DHHS/ASPE recently funded 14 State/County Welfare Leaver Studies. As summarized in ***Exhibit 6.6***, these studies involve the tracking of multiple cohorts of closed TANF cases over varying periods of time using both administrative data and surveys. SSA could contact ASPE to explore what information these studies will provide as currently funded. SSA could also explore working with ASPE and the states to establish SSA data linkages to the cases being tracked and/or to add questions regarding SSI application or receipt among those surveyed. These projects are still in the formative stage; it may be possible for SSA to work with ASPE and the states to make minor changes in data collection plans that would add to the utility of these projects for SSA's purposes.

### Exhibit 6.6 ASPE Welfare Leaver Evaluations

Evaluation	Evaluator	Program Design	Summary
Arizona	None yet	Time limits, family cap, extended transitional Medicaid and child care	The goal of this project is to evaluate (1) whether the sanctions and benefits of the Arizona waiver are successful in motivating participation and employment; (2) whether progressive sanctioning, ending in full-family sanctions motivate employment; and (3) whether families take advantage of the 24-month transitional child care and Medicaid benefits.
Cuyahoga Co., Ohio	Manpower Demonstration Research Corporation and Case Western Reserve University	Comparison of outcomes for Welfare Leavers across two counties in Ohio and California.	Manpower Demonstration Research Corporation (MDRC) will use data from the Urban Change project for a two-site comparison study between Cuyahoga County and the Los Angeles County. The sample includes cohorts from the last calendar quarter of 1996 and first calendar quarter of 1998. The data for the project includes up to ten years of full population administrative data developed for the Urban Change project and mixed-mode sample survey drawn from the second cohort in September 1999.
District of Columbia	The Urban Institute	Little variation from the federal time limits or work requirements.	The Urban Institute will use administrative data from DC's current integrated system to evaluate the impact of welfare reform in DC on individuals whose welfare cases have been closed at least two months. Data are available from 1992 onward. In addition, the study will also include focus groups of individuals who left TANF.
Florida	Florida State will assist with the Survey	Strict time limit (24 or 36 months out of any 60 months, depending on recipient characteristics and previous time on assistance) and generous income disregards in a low-benefit-level state.	This project will address three populations potentially affected by welfare reforms: (1) welfare leavers; (2) those who apply for cash welfare but are never enrolled because of non-financial eligibility requirements or diversion payments; and (3) those who appear eligible but are not enrolled in the state program. Administrative data will be used and telephone surveys of 15,000 households will be conducted over 5 years.
Georgia	Georgia State	4-year time limit, work requirement no later than 24 months after first receiving assistance, family cap, diversion payments of 1-5 months.	The project will build on an on-going study in Georgia by tracking two cohorts of welfare leavers. The first is a cohort of 2,000 leavers will be tracked in administrative records from January to October 1997. A second cohort will track 200 per month from July 1998 to June 2001 via a telephone survey.
Illinois	University of Illinois at Springfield and Chapin Hill	Families with children aged 13 or older have 24-month time limit (otherwise 60-month time limit), family cap, and transition childcare	The University of Illinois at Springfield will build upon an ongoing Closed Case Study using administrative data to track clients for an additional 6 months. A cohort of cases who leave assistance between January and March 1999 will be sampled, using full population administrative data and 800 survey interviews.

**Exhibit 6.6 (Continued)**  
**ASPE Welfare Leaver Evaluations**

Evaluation	Evaluator	Program Design	Summary
<b>Los Angeles County, California</b>	Manpower Demonstration Research Corporation	Comparison of outcomes for Welfare Leavers across two counties in Ohio and California.	MDRC will supplement their Urban Change project to provide a special focus on leavers. The outcomes will be used for a two-site study that allows comparisons between this project and the Cuyahoga County, OH project while controlling for study design. Samples will be taken from cohorts of cases who leave assistance between January and March 1999. Up to ten years of full population administrative data will be used, and mixed-mode sample survey will be drawn from the cohort of cases who leave assistance.
<b>Massachusetts</b>	Chapin Hall and UMASS-Boston	24-month time limit, community service after 2 -months, family cap.	The evaluator will analyze two cohorts of welfare leavers. The first consists of a full population sample (approximately 20,000 cases) of leavers from January to June, 1997, while the second cohort consists population sample (approximately 15,000 cases) of families that exited welfare between December 1998 and February 1999. For the first cohort, data is collected through the MA Dept. Of Revenue's Longitudinal database and a sample survey. For the second cohort, administrative data and a detailed mixed mode survey of 600 cases will be linked to the longitudinal database.
<b>Missouri</b>	University of Missouri and Midwest Research Institute	48-month time limit, extended child care, diversion payments	The Midwest Research Institute will be linking data from state administrative data on health and human services programs, employment and training programs, wage records, and non-profit emergency assistance records to analyze outcomes for welfare leavers.
<b>New York</b>	Rockefeller Institute of Government is advising the project	Diversion payments, expanded earnings disregards, immediate work requirements.	The focus of this project is on several outcomes for welfare leavers including: the frequency of outcomes such as employment, job retention, use of transitional assistance and returns to assistance; identifying barriers to self-sufficiency; examining the effectiveness of sanction policies in changing behavior; and developing a longitudinal tracking capacity for welfare outcomes in New York City.
<b>San Mateo, Santa Clara and Santa Cruz Counties, California</b>	Sphere Institute and TBD	Comparison of outcomes for welfare leavers across three counties in California.	The SPHERE Institute plans to work with a consortium of three contiguous counties (San Mateo, Santa Cruz, and Santa Clara) to evaluate the impacts of welfare reform on policy relevant subgroups in these counties in 1996 and 1998. Administrative data will be used in an analysis of cohorts in 1996 and 1998. In addition, two thirty- minute mixed mode surveys will be conducted for the 1998 cohort.
<b>South Carolina</b>	Under negotiation	24 month time limits out of 120 months, subsidized employment, no transitional Medicaid longer than 12 months.	The evaluator will analyze individuals who have left welfare and stopped receiving benefits for four months. Two cohorts will of job losers will be used from 1997 and 1999. Outcomes of interest include changes in marital status, employment, and earnings.



**Exhibit 6.6 (Continued)**  
**ASPE Welfare Leaver Evaluations**

Evaluation	Evaluator	Program Design	Summary
Washington	None yet	Transition child care if income does not exceed 175% of federal poverty level, \$1500 diversion payment limit.	The focus of this project is on three populations: those who receive welfare; those who are diverted, and those who appear to be eligible but do not enroll. The sample for the analysis includes three cohorts: a pre-TANF cohort from the fourth quarter of calendar year 1996; an early implementation cohort from the fourth quarter of 1997; and a full implementation cohort from the fourth quarter of 1998. The second cohort includes a sample of continuing cases for comparison, while a mixed mode survey of 1,300 cases is planned for the third cohort. For all cohorts, linked administrative data will be used from TANF Food Stamps, Medicaid, Child support, Child Welfare, Unemployment Insurance and the State Basic Health plan for the 24 months around exit time.
Wisconsin	None yet	Diversion strategy, strict work requirements	This study will expand upon three existing projects. The first project will develop a longitudinal database from 1998 forward to study families who left AFDC prior to the implementation of Wisconsin Works (W-2) or who did not convert during the transition. The second project will include a survey (that is already in the field) for clients who leave W-2 in 1998. The final project will expand a planned study of people who apply for W-2 in Milwaukee between October 1998 and March 1999.

SSA could also explore further the extent to which the Urban Institute’s “Assessing the New Federalism Project” addresses issues of interest to SSA.<sup>330</sup> At a minimum, the study will provide detailed information on state policies in all states, case studies of program implementation in thirteen states, and information on the status of low-income families in those 13 states. Six of the Urban Institute states (California, Florida, Massachusetts, New York, Washington, and Wisconsin) are also participating in the ASPE Welfare Leavers Study. The Urban Institute’s case study findings from these six states could nicely complement the tracking information obtained through the welfare leavers study. It might be possible to explore the addition of questions of special interest to SSA such as the treatment of persons with disabilities or the active referral of TANF clients to SSA to the current case study protocol. In **Exhibit 6.7** we list the states identified as participating in experimental studies of interest, welfare leavers projects, and/or the Urban Institute Study.

**Exhibit 6.7**  
**Summary of Evaluations by State**

State	Experimental Evaluation	Welfare Leaver Project	New Federalism State <sup>331</sup>
Alabama			X
Arizona	X	X	
California	X <sup>332</sup>	X <sup>333</sup>	X
Colorado			X
Connecticut	X		
District of Columbia		X	
Florida	X	X	X
Georgia		X	
Illinois		X	
Indiana	X	X	
Iowa	X		
Massachusetts		X	X
Michigan			X
Minnesota	X		X
Mississippi			X
Missouri		X	
Nebraska	X		
New Jersey			X
New York		X	X
Ohio		X <sup>334</sup>	
South Carolina		X	
Texas	X		X
Vermont	X		
Washington		X	X
Wisconsin	X	X	X

<sup>330</sup> Urban Institute’s “Assessing the New Federalism Project” is available at [www.newfederalism.urban.org](http://www.newfederalism.urban.org).

<sup>331</sup> Represents one of the “focus” states in The Urban Institute’s ANF project.

<sup>332</sup> Experimental design is included in several county evaluations.

<sup>333</sup> Several California counties have welfare leaver projects.

<sup>334</sup> Cuyahoga County in Ohio has a welfare leaver project.

SSA could also undertake its own tracking studies and case studies. SSA may want to identify a sample of states of special interest, and set up data matching arrangements with those states to track transitions from TANF to SSI over time. Such arrangements could build on and supplement tracking data collected through the ASPE studies by tracking a larger sample of TANF families for a longer period of time than anticipated in those studies. We suggest exploring this possibility initially in Florida and California. Over time, SSA could expand tracking to other large states.

SSA case studies in these same states could provide more SSA-specific information on the aspects of TANF implementation that are of particular interest to SSA as well as the perceptions and experiences of personnel in SSA field offices and state disability determination offices. For example, SSA could explore whether there are specific state efforts to divert TANF applicants or recipients into SSI, and, if so, how they are structured. SSA could also explore state efforts to provide employment and training services for people with disabilities who are not currently receiving SSI. State success in this area could help prevent future SSI applications and may have relevance to SSA's own work initiatives for people with disabilities.

All of these descriptive study approaches will provide SSA information on the flow of TANF recipients into SSI and on implementation choices states are making that may be influencing those transitions. If collected over time in a number of states, this information may be used to support future modeling efforts of the effects of TANF on SSI. At a minimum, it will enable SSA to place the knowledge gained through the experimental studies in a larger context.

It would be especially useful to conduct case studies in states in which evaluations of interest to SSA are being conducted, and in states that account for large shares of the SSI caseload. There are, however, operational criteria that must be considered when states are selected for this purpose. Specifically: key individuals in the state (managers, administrators and technicians) must be willing to invest necessary time and effort; operating systems, operating procedures, and state personnel must be able to provide the necessary information; and the work must have sufficient priority to successfully compete with other state activities for scarce resources.

### **C. DID Estimates**

Linking either the Welfare Leavers data or the state administrative data to SSA data would offer an opportunity to validate DID analyses of SSI outcomes that use the SSA administrative data alone. The state administrative data might be especially useful because they would allow assessment of transitions from AFDC to SSI prior to TANF. As a first step, these data could simply be used to verify the accuracy of the SSA administrative data concerning AFDC/TANF participation at the time of application. Beyond this, however, the data could be used to identify SSI applicants who were not AFDC/TANF recipients at the time of SSI application, but who were former AFDC/TANF recipients. These could then be included in the target group for the analysis of the impact of TANF, rather than in the comparison group. If the number of such recipients is large, then the results might be quite different than those obtained from the administrative data alone.

While these estimates will likely be of better quality than those based on SSA administrative data alone, they will nonetheless be subject to an important *caveat*: the estimates are only as good as

the assumption that other factors affecting SSI outcomes for the target group have proportional impacts on outcomes for the comparison group. Hence, the estimates would not be as strong as those that might be obtained from the experimental evaluations described in the previous section.

## VII. OTHER DATA SOURCES

We have collected and assessed information on a variety of data sources that would be potentially useful for the evaluations. The most promising of this is The Urban Institute's Welfare Rules Database, which includes detailed information on state TANF programs. SSA may want to influence information that Urban is collecting concerning TANF recipients with disabilities. Several administrative data sources other than those previously mentioned are also hold promise. They might be linked to SSA data and used for analysis similar to that described in the previous section for the case studies. National surveys other than the SIPP and CPS show little promise of value to the evaluation.

### A. Information on State TANF Programs

One source that may be very valuable for making cross-state comparisons of the effects of TANF on SSI is The Urban Institute's Welfare Rules Database. The database includes detailed information on several aspects of individual state TANF programs. Of particular interest are variables being collected on special provisions for persons with disabilities. Presumably, the effects of TANF on SSI would be larger in states that have fewer exemptions from work requirements or time limits for persons with disabilities. In *Exhibit 6.8*, we highlight some of the questions that are being addressed by The Urban Institute that deal directly with treatment of persons with disabilities. This database also contains several other questions that could be of use in a pooled analysis (e.g., employment-related rules).

#### Exhibit 6.8

#### Questions from the Urban Institute's Welfare Rules Database Regarding TANF Provisions for Persons with Disabilities

##### State TANF Eligibility

- Are SSI recipients eligible for benefits?

##### Work Requirements

- Are ill or incapacitated persons exempt?
- Are persons caring for an ill or incapacitated member exempt?

##### Time Limit Requirements

- Are ill or incapacitated persons exempt?
- Are persons caring for an ill or incapacitated member exempt?

SSA may want to investigate whether some of the questions can be expanded to explore specific transitions from AFDC to SSI. For example, the database currently includes questions on whether a state provides assistance in the form of a one-time cash payment, support services, or both to divert applicants or recipients from the state TANF program. This question could be expanded to ask whether a state has specific policies (formal or informal) to divert TANF recipients to SSI, such as requiring persons with disabilities to apply for SSI while they receive

TANF benefits.<sup>335</sup> Additional questions could also be added for the treatment of SSI income for both children and adults in calculating TANF benefits.<sup>336</sup> Such information would provide some indication of each state's effort in trying to divert potential applicants and recipients from AFDC to SSI.

## B. Other Administrative Data Sources

We identified several administrative data sources that could be useful for SSA's purposes that were not included in our welfare reform evaluation review (Chapter 2) or site visit discussion (Chapter 3). The first administrative data source is the Integrated Database on Children's Services (IDB) in Illinois. The IDB contains longitudinal records on any child that was in contact with any of the following state services: foster care, child abuse, special education, mental health, juvenile justice, Medicaid, Food Stamps and AFDC. These data also include SSNs that could be used to link the IDB to SSA data. The merged data could be used to analyze transitions from several state programs to SSI in Illinois over several years. According to Goerge, *et. al.* (1996) these data also include information on SSI receipt in years following 1994.<sup>337</sup> They used these data to identify service utilization and the characteristics of children with disabilities in Illinois from 1990 to 1994. Based on their tabulations in 1994 alone, there were 277,689 disabled children in their database.<sup>338</sup> Hence, this database should be sufficiently large to analyze transitions from AFDC to SSI.

A second potential administrative data source is being constructed by The MEDSTAT Group for the Substance Abuse and Mental Health Services Administration (SAMHSA). MEDSTAT is developing national spending estimates for mental health and chemical dependency services, using an integrated database for three state mental health, chemical dependency, and Medicaid programs. These data may be of some use for SSA's purposes, though it could be difficult to use Medicaid data alone to identify transitions from state programs to SSI.<sup>339</sup> We believe the administrative data sources identified earlier in this report would be better suited for SSA interests.

In *Lewin (1998b)*, we identified two state administrative data sources that could potentially be linked to SSA records. The first is being constructed in Missouri. The state of Missouri has been collecting data for adults receiving AFDC payments along with information on wage records from the state's Unemployment Insurance program since 1992. These data have been made available to researchers at the University of Missouri-Columbia. Dr. Kenneth Troske at

<sup>335</sup> During our site visit in Connecticut, one of the first questions asked was whether a person had a disability. If so, they were directly referred to the SSA office.

<sup>336</sup> Currently the database contains questions for certain types of income (e.g., dividend income, Earned Income Tax Credit).

<sup>337</sup> In their report, Goerge, et.al. did not have access to records on SSI receipt. As a proxy for SSI receipt, they use records for individuals who received Assistance for the Aged, Blind and Disability programs (AABD) from 1990 to 1994. They find that AABD is an excellent proxy for SSI receipt based on administrative records.

<sup>338</sup> They identified children with disabilities based on program participation in special education, AABD, mental health services, or Medicaid (for those who received reimbursed service for preventative, well-child care, and more serious inpatient rehabilitative services).

<sup>339</sup> For example, in some states individuals are not categorically eligible for Medicaid through their TANF participation. Hence, the state may not track other state program participation in its Medicaid population.

the University of Missouri-Columbia is planning to use these data to track transitions from welfare to work. The second administrative data source is being constructed in Maryland. Dr. Catherine Born used an administrative database with wage/employment files and interviews to track a representative sample of over 2,000 families who left the welfare system during the period in which the Maryland Family Investment Program was being implemented (October 1996 to September 1997). In addition to these data, Dr. Born has used state administrative data from previous years to study a cohort of Maryland welfare families to analyze exits from welfare rolls (Caudill and Born, 1997). Dr. Born indicated that the state government of Maryland was very cooperative in assisting her evaluation efforts and noted that they are an excellent candidate to link SSA data to state data. While both of these data sources are potential candidates for data linking, the state administrative data sources identified in the welfare reform evaluation reviews are likely to be more promising because of their experimental and/or state TANF program design.

### C. Other Survey Data Sources

In addition to the SIPP and Survey of American Families, there are other potential survey data sources that SSA may want to consider in a future welfare reform evaluation. The first is the Survey of Program Dynamics (SPD) that is being put together by the Census Bureau. The SPD uses an overlapping sample from the 1992 and 1993 SIPP panels and follows them for six years from 1996 to 2001.<sup>340</sup> The SPD could be linked with the 1992 and 1993 SIPP panels to create a ten year panel data set of employment and program participation from 1992 to 2001.<sup>341</sup> These data could be used to analyze transitions from state programs to SSI over the period of the welfare reform changes. Another advantage of using the SPD is that SSA has already linked administrative records to the 1992 and 1993 SIPP panels. Hence, all of the individuals in the SPD will have SSA information from SSA data sources.

One potentially major drawback of the SPD is attrition bias. In Chapter 5, we found that attrition rates were higher for SSI applicants and recipients than others. Because the SPD target sample includes only individuals who responded to the final SIPP interview, there will be some selection bias in the initial sample. Our finding leads us to believe that SSI applicants and recipients will be underrepresented. Further, according to Huggins and King (1998), the sample attrition after the last SIPP interview for the 1998 SPD was very high -- 50 percent. Evaluations that rely on the SPD to evaluate welfare reform will need to account for these attrition biases.

A second potential data source is the Panel Study of Income Dynamics (PSID). The PSID is a longitudinal file that contains detailed demographic, health, program, and income information for a nationally representative sample. Currently, the PSID contains longitudinal data from 1968 to 1995. While these data could potentially be used in a future evaluation option to analyze long term transitions, we believe the SIPP and CPS provide more viable options because of their sample size (a typical PSID cross-section has approximately less than half the number of observations than a typical SIPP panel) and the availability of matched SSA data.

---

<sup>340</sup> Individuals who completed both the first and last wave of the 1992 and 1993 Panels are included in the SPD target sample.

<sup>341</sup> The first sample in the 1992 SIPP panel was interviewed in February of 1992.

A third potential data source is the American Community Survey (ACS).<sup>342</sup> The Bureau of the Census is developing the ACS as a tool for collecting data at the community level that are currently collected only in decennial census years, via the census long form. Thus, the ACS offers the potential opportunity to produce state and local estimates that cannot be supported by the CPS or, for that matter, SIPP.

In comparison to the CPS and SIPP, the ACS data are much less detailed and comprehensive. For instance, the current version of the instrument does not distinguish between TANF, SSI and general assistance income, nor does it identify the individuals in the household who are the recipients. Identification of SSI recipients might be accomplished by linking the ACS data to SSI administrative data. Linked data would allow SSA to track SSI applications, allowances and caseloads by demographic group for states and major metropolitan areas (potentially all areas with populations of at least 65,000 persons or more) annually. Thus, for instance, reasonably accurate annual estimates of the percent of young women with children who apply for, are awarded, or receive SSI in each year could be produced for each state, and compared to the same proportion for young women without children.

The ACS could be useful for monitoring program interactions in the future, but by itself will not be very useful for evaluating the recent reforms because it will not be fully implemented until 2003. Estimates for states and all area with populations of at least 250 thousand are planned for 2001. One could also use the 1990 and 2000 Census long form samples to obtain estimates for those years. The desirability of conducting analyses using data from these combined sources is, however, reduced by comparability problems and lack of intermediate year data. It would be especially problematic to isolate the effects of the reforms from the many other policy and environmental changes that occurred between 1990 and 2000. The logistical challenge of matching the ACS to SSA administrative data might also be a significant deterrent to this activity.

SSA is funding two surveys that will eventually yield substantial information about interactions between SSA programs and other programs. The first of these is a survey of children who were potentially affected by the child SSI reforms, including a sample of those who were not already SSI recipients when the legislation was passed. The second is the Disability Examination Study (DES), which will examine a nationally representative, stratified sample of approximately 5,000 working age individuals with severe disabilities. Most DES respondents will not be SSI or SSDI beneficiaries. A substantial number are likely to be current or recent TANF recipients.

The DES will assess whether each examined respondent meets the medical eligibility criteria for SSI and SSDI, and will produce estimates of the prevalence of disability, defined by these criteria, in the general population. Thus, the DES should provide estimates of both the number of adult TANF recipients and the number of low-income adults with children who are at high risk for SSI. This by itself will not be directly helpful in assessing the impact of the recent reforms because DES data collection is not scheduled for completion until 2001. The DES data will, however, offer an opportunity to develop prediction models for SSI eligibility from SIPP,

---

<sup>342</sup> Information on the ACS contained in this discussion was obtained from the Bureau of the Census web site, [www.census.gov/acs](http://www.census.gov/acs).



CPS and possibly other national surveys. A SIPP based model could eventually be used in analysis of the matched SIPP/SSA data.

One final survey worth brief mention is the National Health Interview Survey, conducted by the Department of Health and Human Services. This periodic survey collects extensive health and disability data on a very large, nationally representative sample, but usually collects little information on income and program participation. Our understanding is that these data cannot be linked to SSA data due to confidentiality issues and lack of necessary identifiers. Hence, these data would not be very useful in an evaluation of the impact of the recent reforms on SSI.

It is unrealistic to expect accurate national estimates of the total impacts of all reforms, or of specific non-SSA reforms alone. There is, however, much that can be done to obtain useful information about the interactions between SSA and non-SSA programs, the intersection between the populations they serve, and how they both are changing over time because of program changes as well as other factors. SSA administrative data and matched Census/SSA data offer opportunities to conduct these types of analyses.

We have identified a set of complementary approaches for expanding SSA's understanding of the effects of non-SSA reforms on SSA programs (*Exhibit 6.9*). The best way to rigorously evaluate the impacts of non-SSA reforms on SSI is by building on experimental welfare evaluations currently underway. Even though these evaluations will not produce nationwide estimates of the impacts of reforms, they offer a unique opportunity to establish a causal relationship between specific TANF reforms and SSI outcomes. SSA can supplement information it gathers through experimental studies by conducting case studies of specific states using both quantitative and qualitative methods. SSA can use state administrative data, SSI administrative data, and survey research to track the SSI applications of current and former TANF recipients and the SSI allowances of former TANF recipients. Qualitative case study data can be used to provide contextual information regarding the TANF program and policy initiatives that influence the movement of clients from TANF to SSI.

The analysis of SSA administrative data provides impact estimates for every state. These can be validated in at least some states. The opportunity to validate is strongest in states that are conducting experimental welfare reform evaluations. Pursuit of the latter can be expected to produce quite definitive evidence of the impacts of reforms in these states. In other states validation is limited to using state administrative data or survey data to verify SSA data and to make marginal improvements in the analysis that can be conducted with the SSA data alone. The analysis of matched Census/SSA data complements all of these activities because it offers the opportunity to follow transitions to SSA for samples that are representative of the national population. This analysis can take advantage of the rich individual information that is available in SIPP or the CPS, but linking changes in transitions to SSI to the state reforms will be problematic because of the small samples of transitions observed in each state.

The options for evaluating the combined impacts of SSA and non-SSA reforms are more limited. The analysis of SSA administrative data can produce state and national estimates of the combined effects of the TANF, DA&A, and child SSI reforms on applications, allowances, caseloads and benefits for those who were not SSI recipients or DI beneficiaries at the time the

reform legislation was enacted. While these cannot be validated in ways that are comparable to the validation opportunities available for estimating the impacts of TANF, what is learned from validating the latter would be useful in interpreting the estimates of the combined reforms. The analysis of the matched CPS/SSA data can provide further information on caseload impacts, including impacts on pre-reform recipients. The information gained from these analyses would supplement the findings from the separate evaluations of the DA&A and child SSI reforms. First-cut estimates of the impacts of non-citizen reforms can be obtained through analysis of the administrative SSI.

**Exhibit 6.9**  
**Summary of Evaluation Options**

Options		Reforms			Outcome Variables			
		AFDC/ TANF	Non-Citizens	Total	Applications	Allowances	Caseload	Benefits
1. SSA Administrative Data	DID*	0	0	0	0	0	0	0
	Pooled Time Series	0		0	0	0		
2. Matched Census/SSA Data	Hazard Analysis**	0			0	0	0	0
	Caseload Analysis	0		0			0	0
3. Welfare Impact Evaluation Add-ons***		0			0	0	0	0
4. State Case Studies***		0	0		0	0	0	0

\*Difference in Differences analysis.

\*\*Includes auxiliary analysis of benefit continuation and payments for allowed applicants.

\*\*\* In selected states only.